Comparison of Magnetic Resonance Spectroscopy (MRS) data in children with and without HIV at 11-12 years

Amy Graham

(GRHAMY001)

Submitted to the University of Cape Town in fulfilment of the requirements for the degree

MSc(Med) in Neuroscience

Supervisors: Dr Frances Robertson and Professor Ernesta Meintjes

Division of Biomedical Engineering

Department of Human Biology

Faculty of Health Sciences

University of Cape Town

10 September 2019

Declaration

I, Amy Graham, hereby declare that the work on which this dissertation is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: 10/9/19

Acknowledgements

There are several people who I would like to acknowledge and thank for making this Master's project a reality. Thank you to my supervisors, Dr Frances Robertson and Professor Ernesta Meintjes, for your continuous guidance and support throughout my Master's degree. I am very grateful to have had the opportunity to work under your supervision and for how you have encouraged me to learn new skills and try out different analytical approaches. Thank you too, Dr Robertson for all the coding that you did for the processing of the spectra and for all your patience in helping me with the data processing and analysis.

Thank you to the National Research Foundation (NRF) of South Africa and the University of Cape Town for providing funding for my Master's studies (DST-NRF Innovation Masters scholarship, Grant Number: 117183; UCT Conference Travel Grant; Masters VC Research Scholarship). This project was funded by the NRF/DST South African Research Chairs Initiative; NRF grant CPR20110614000019421; South African Medical Research Council (SAMRC); US National Institute of Allergy and Infectious Diseases (NIAID) (CIPRA network, Grant U19 AI53217); and NIH grants (R01HD085813, R01HD071664, R01HD093578, R21MH096559 and R21MH108346).

To the various collaborators who have contributed to this work, I am grateful for all your input and suggestions for conference poster presentations and as I have prepared a manuscript. Thank you to Dr Holmes in the Division of Biomedical Engineering and Professor Little in the Department of Statistical Sciences at the University of Cape Town, for all your guidance in this project. Thank you, Dr Laughton, Professor Cotton and Dr Dobbels, at Stellenbosch University, for the clinical data which you provided. Thank you, Professor van der Kouwe at Harvard Medical School for the scanning sequences which you provided for this work. Thank you too to the radiographers and research staff at CUBIC for all the scanning carried out to supply the spectra for this study.

Thank you to all the members of my lab who have given me advice and help along the way. I appreciate the friendships we have built over these two years. Thank you Dr Warton for suggesting that I consider doing a Master's in Neuroscience. I am so grateful for all your advice. Thank you too to the Andersons and other friends for supporting me along the way.

Finally, I am so grateful to my family who have cheered me on and encouraged me throughout this project. Thank you for your continuous love and support.

Soli Deo Gloria. Romans 11:36 'For everything comes from him and exists by his power and is intended for his glory. All glory to him forever! Amen.'

Research outputs

Aspects from this project have been presented at the following conferences/workshops:

- The Second HIV and Adolescence workshop in Cape Town, South Africa, which took place on 10-12 October 2018 – a poster was presented.
- The International Society for Magnetic Resonance in Medicine (ISMRM) 27th Annual Meeting in Montreal, Canada, which took place on 11-16 May 2019 – a digital poster was presented.
- The Organisation for Human Brain Mapping (OHBM) Annual Meeting in Rome, Italy, which took place on 9-13 June 2019 – a poster was submitted.

Abstract

Although HIV and antiretroviral drugs have been shown to cause damage in the brain, the long-term impacts of perinatal infection, early treatment and exposure in children at 11 years, remain unclear. The effects of HIV and antiretroviral therapy (ART), whilst indistinguishable, can be investigated at a chemical level through proton magnetic resonance spectroscopy (1H-MRS).

Previous studies in children have largely focused on individual metabolite changes. However, several adult studies have now advanced beyond this to address patterns of metabolic activity that are altered with HIV infection.

Using a 3T Skyra scanner, 136 children (76 HIV+, 30 HEU, 30 HU; 71 males) between the ages of 11.0-12.5 years, and from a similar socioeconomic background, were scanned. In this study metabolite concentrations were quantified within the basal ganglia (BG), midfrontal gray matter (MFGM) and peritrigonal white matter (PWM). We utilised linear regression to investigate individual metabolite differences, comparing HIV-infected (HIV+) children from the Children with HIV Early Antiretroviral Therapy (CHER) trial, and HIV-exposed-uninfected (HEU) children, to HIV-unexposed (HU) children. Pearson's correlation analysis, factor analysis and logistic regression were then used to study alterations in metabolic patterns between HIV+ and HIV-uninfected (HIV-) children. Analysis of the data was carried out in R.

We found elevated total choline in the BG ($p = 0.03$) and MFGM ($p < 0.001$) of HIV+ children, as well as reduced PWM total NAA ($p = 0.03$) and total creatine ($p = 0.01$). Altered metabolite concentrations were further observed in HEU children. Additionally, we identified a cross-regional coupling of choline which distinguishes HIV+ from HIV- children ($p < 0.001$).

These findings indicate that multiregional inflammation and PWM axonal damage are occurring in HIV+ children at 11 years. Ultimately, the consequences of perinatal HIV acquisition, in spite of early treatment, continue to be seen at 11 years, as do the impacts of exposure.

Table of contents

List of tables

List of figures

Abbreviations

AIDS: Acquired immune deficiency syndrome
ADC: AIDS dementia complex.
ANOVA: Analysis of variance.
ART: Antiretroviral therapy.
ART-Def: The deferred ART group.
ART-40w: Group on early ART for 40 weeks.
ART-96w: Group on early ART for 96 weeks.
ARV: Antiretroviral.
ATP: Adenosine triphosphate.
AUC: Area under the curve.
BBB: Blood brain barrier
BG: Basal ganglia.
BOLD: Blood oxygenation level dependent
cART: Combination antiretroviral therapy.
CDC: Centers for Disease Control
CHER: Children with HIV Early Antiretroviral therapy trial
CHESS: Chemical shift specific.
Cho: Choline.
CNS: Central nervous system.
Cr: Creatine.
CSF: Cerebrospinal fluid.
DMN: Default mode network.
FA: Fractional anisotropy.
FC: Functional connectivity.
FWHM: Full width at half maximum.
Glu: Glutamate.
Glx: Glutamate + glutamine.
GM: Gray matter.
GMDS: Griffiths Mental Development scales.
HEU: HIV-exposed-uninfected individuals.
HIV: Human Immunodeficiency Virus.
HIVE: HIV encephalopathy.
HIV+: HIV-infected.
HIV-: HIV-uninfected.
HU: HIV-unexposed individuals.
Ins: Myo-inositol.
IQR: Interquartile range.
LCModel: Linear Combination Model software.
MDM: Magnetic dipole moment.
MEMPRAGE: Multi-echo magnetisation prepared rapid gradient echo.
MFGM: Midfrontal gray matter.
MRI: Magnetic Resonance Imaging.
MRS: Magnetic Resonance Spectroscopy.
NAA: N-acetyl-aspartate.
PMTCT: Prevention of mother-to-child transmission.
PRESS: Point Resolved Spectroscopy.
PWM: Peritrigonal white matter.
RF: Radiofrequency.
ROC: Receiver operating characteristic.
SNR: Signal-to-noise ratio.

SPM12: Statistical Parametric Mapping software.
STEAM: Stimulated Echo Acquisition Mode.
SVS: Single voxel spectroscopy.
tCho: Total choline (glycerophosphocholine + phosphocholine)
tCr: Total creatine (creatine + phosphocreatine)
TE: Echo time.
tNAA: Total NAA (N-acetyl-aspartate + N-acetyl-aspartyl-glutamate)
TR: Repeat time.
VL: Viral load.
WHO: World Health Organisation.
WM: White matter.
1H-MRS: Proton Magnetic Resonance Spectroscopy.

1. Introduction

New HIV infections continue to occur in children despite the advancement in facilitating better access to antiretroviral therapy (ART) for pregnant HIV-infected (HIV+) women across the globe (UNAIDS, 2019). This may be due to various factors including challenges of adherence, premature births and incomplete protection provided by ART (Stringer et al., 2010; Warszawski et al., 2008).

In order to determine when to initiate treatment in seropositive infants in South Africa, a randomised clinical trial was established - the Children with HIV Early Antiretroviral Therapy (CHER) trial (Cotton et al., 2013; Violari et al., 2008). This trial confirmed the value of early treatment in infants under 12 weeks of age, reducing mortality rates and morbidity, as seen by improved short-term cognition and health status (Cotton et al., 2013; Laughton et al., 2012; Violari et al., 2008).

Adherence to ART regimens is essential, as the interruption of treatment can enable the re-emergence of infection due to persisting viral reservoirs (Joos et al., 2008). However, studies in animals have shown detrimental neurological effects that result directly from antiretroviral drugs (Robertson, Liner & Meeker, 2012) and, moreover, HIV is associated with functional and structural abnormalities in HIV+ children on ART, despite viral suppression (Toich et al., 2018; Laughton et al., 2018; Randall et al., 2017). Therefore, this raises questions regarding the long-term implications of treatment.

It should be noted that brain development is not only impacted by HIV and ART. There are other contributing factors, such as environmental exposures and events, that influence brain development in children (Paul et al., 2019; Walker et al., 2007) and that have been shown to play a role in cognitive and behavioral changes that occur in HIV+ children (Debeaudrap et al., 2018; Mellins et al., 2003). These must also be kept in mind when carrying out studies of HIV+ children on treatment.

With the introduction of effective ART for pregnant HIV+ women, leading to a reduction in mother-to-child transmission of HIV (UNAIDS, 2019; De Cock et al., 2000), the population of perinatally HIV-exposed-uninfected (HEU) children has increased - with possible mental health and attention issues among these children (Kerr et al., 2014; Malee et al., 2011).

Thus, there are two populations of interest, HIV+ children on ART and HEU children, who may have been exposed to ART pre- and/or post-natally. When studying the impacts of perinatal infection and exposure, the effects of ART drugs cannot be isolated and must be taken into consideration.

Magnetic resonance imaging provides a technique for investigating changes that occur in the brain as a result of HIV infection and early ART, or due to HIV exposure. Single voxel proton magnetic resonance

spectroscopy (1H-MRS) specifically examines the underlying metabolic changes in select brain regions, in a non-invasive way (Wilkinson et al., 1997).

Adding to previous cross-sectional evaluations of HIV+ children from the CHER trial, who were placed on the same ART regimen before the age of 2 years, this study seeks to examine whether the effects of HIV and ART persist, specifically as the children enter into adolescence, or if metabolic activity is restored. We endeavor to examine differences in metabolite concentrations between HIV+ or HEU children and HIV-unexposed (HU) children, in three brain regions of interest – the basal ganglia (BG), peritrigonal white matter (PWM) and midfrontal gray matter (MFGM) – in a subset of children from the CHER trial and in HEU children at the age of 11 years. Additionally, we aim to investigate how inter-regional and inter-metabolite relationships (as in Yiannoutsos et al., 2004) are impacted by HIV infection and treatment.

2. Literature review

2.1 Introduction to HIV

2.1.1 The challenges of targeting HIV

The Human Immunodeficiency Virus (HIV) is responsible for a global pandemic which began during the 1980s (Sepkowitz, 2001) and remains a major issue today. HIV is an effective pathogen due to its ability to attack and evade the host defense system (Allen et al., 2005; Siliciano et al., 2003; Nowak & McMichael, 1995). The $CD4^+$ T cells, necessary for responding to foreign bodies, fighting infection and aiding other components of the immune system, are the main target cell population within a host, together with other CD4 receptor-expressing cells (Alimonti, Ball & Fowke, 2003; Berger, Murphy & Farber, 1999). Irreversible damage to the host immune system can occur, particularly in infants, within the first few months of HIV infection (Derdeyn & Silvestri, 2005; Correa & Munoz-Fernandez, 2001; Kourtis et al., 1996). Additionally, a progression of HIV infection in individuals who do not receive treatment leads to acquired immune deficiency syndrome (AIDS). At this stage of disease, individuals are highly susceptible to other opportunistic infections (Alimonti, Ball & Fowke, 2003; Weiss, 1993).

HIV establishes latent reservoirs early during infection, with the Central Nervous System (CNS) as a key target site due to the low accessibility of antiretroviral (ARV) drugs to this system (Thompson et al., 2011; Palmer, Josefsson & Coffin, 2011; Varatharajan & Thomas, 2009). Viral reservoirs enable the pathogens to persist within their host, leading to a resurgence of infection if treatment is discontinued (for review see Dahl, Josefsson & Palmer, 2010; Joos et al., 2008). Antiretroviral therapy (ART) is therefore a life-long commitment in HIV-infected (HIV+) individuals.

2.1.2 Antiretroviral therapy

The introduction of combination ART (cART) has resulted in a significant reduction in HIV-related death and severe disease (Robertson et al., 2007; Brodt et al., 1997), including opportunistic infections and encephalopathy (Martin et al., 2006). cART can target multiple points in the viral life cycle simultaneously, thus making it more difficult for the virus to develop mutations to escape the drugs (Kurth et al., 2011). However, despite the use of cART, HIV is still capable of mutating to develop resistance to drug challenges, leading to the failure of specific ART regimens (Hamers et al., 2012). As a result, this may impact which drugs can be used in further treatment (Hamers et al., 2012) and patients may need to change to second-line regimens to target other aspects of the viral life-cycle.

Even in patients who adhere well to treatment, HIV is able to develop drug resistance (Bangsberg et al., 2003). However, it has been found that poor or varying adherence to ART favours the survival of

virions with mutations that relay resistance (Gardner et al., 2008; Clavel & Hance, 2004). HIV is a retrovirus which lacks a means of proofreading while replicating the single-stranded RNA which constitutes its genetic material, unlike DNA replication mechanisms (Jonckheere, Anné & De Clercq, 2000). The rapid replacement of virions (Perelson et al., 1996) and a high mutation rate (Sanjuán et al., 2010) better enable HIV to rapidly achieve drug resistance. The ability of HIV to develop resistance has been found to differ for different classes of antiretroviral drugs, affording HIV the ability to overcome these drugs (Gardner et al., 2009; Clavel & Hance, 2004). For example, mutations can prevent protease inhibitors from recognizing the HIV protease enzyme and carrying out their roles or can enable HIV to remove chain-terminating nucleosides that have been incorporated by nucleoside reverse transcriptase inhibitors during replication (Clavel & Hance, 2004).

As of 2018, 7.7 million people in South Africa were estimated to be living with HIV, of whom approximately 260,000 were children (UNAIDS, 2019). There has been a reduction in vertical HIV transmission from mothers to children, with the number of new HIV cases in children below the age of 15 dropping from 28,000 in 2010 to 14,000 in 2018 (UNAIDS, 2019). This reduction may largely be credited to the increased accessibility of ART among pregnant HIV+ women. In 2018, 87% of pregnant women with HIV in South Africa had access to ART (UNAIDS, 2019). As the number of cases of vertical HIV transmission drop (UNAIDS, 2019; De Cock et al., 2000), there is a growing population of HIV-exposed-uninfected (HEU) children.

However, despite the introduction of initiatives to prevent mother-to-child transmission (MTCT), new cases of HIV in infants still occur (UNAIDS, 2019; Warszawski et al., 2008), making the early treatment of HIV+ children a relevant issue particularly in South Africa. Limited resources and the late diagnosis of pregnant mothers are partly responsible for the continuation of perinatal HIV infections, due to not all pregnant HIV+ women receiving ART (Nkonki et al., 2007). However, poor adherence to ART has also been suggested to play a role in new infections occurring in infants as maternal viral levels are less effectively suppressed (Myer et al., 2017).

Research assessing the optimal time for commencing treatment has led to amendments in the World Health Organisation (WHO) recommendations for the initiation of ART. In 2013, early ART was conditionally recommended for HIV+ infants (WHO, 2013). The most recent revision by the WHO in 2017 recommended prioritizing the immediate initiation of treatment in children below the age of 5 years, following a positive HIV diagnosis (WHO, 2017). This advice was based on evidence from various randomized trials, such as the CHER trial (Violari et al., 2008) which will be discussed shortly.

2.2 HIV infection of the brain

2.2.1 Neurological targets of HIV

Studies addressing the effects of HIV on the brain have led to a better understanding of how HIV can infect the brain, establish reservoirs and cause pathology. During HIV infection, the integrity of the blood brain barrier (BBB) is compromised (Lamers et al., 2011), with viral proteins such as Tat playing an important role in this process and driving a strong host inflammatory response . Disruption of the blood-cerebrospinal fluid (CSF) barrier also occurs and, thus, HIV RNA can be detected in the CSF through lumbar puncture during acute infection (de Almeida et al., 2016; Spudich et al., 2011). However, the immune response mounted in the CSF is differentially characterized from that of the plasma and has a greater magnitude during HIV infection (de Almeida et al., 2016).

Once in the brain, HIV does not directly damage neurons, but rather targets supporting cells such as microglia and astrocytes (for review see Churchill & Nath, 2013; Anthony & Bell, 2008; Anderson et al., 2002; Kaul, Garden & Lipton, 2001). Microglia express the CCR5 co-receptor which, together with the CD4 receptor, is required for HIV to initially gain entry into the cells (Vallat et al., 1998).

The main roles of the supporting cells include releasing signaling molecules, such as cytokines and growth factors, which favour neuronal growth (Chen & Trapp, 2016; Thored et al., 2009; Anthony & Bell, 2008) and play a role in regulating the immune response, largely by opposing proinflammatory responses (Chen & Trapp, 2016; Tarassishin, Suh & Lee, 2011). However, once infected with HIV, microglia release viral proteins, pro-inflammatory cytokines and other neurotoxic substances (Anderson et al., 2002).

It has been suggested that the release of excess glutamate from microglia and macrophages is largely responsible for damage to neurons, and may be a good target for treatment strategies addressing HIV-associated neurological disorders (Jiang et al., 2001). High levels of glutamate act via N-methyl-D-aspartate receptors to alter ion transport in neurons, disrupting signal transduction and placing the neurons under strong osmotic stress (Barger et al., 2007; Takeuchi et al., 2005). Thus, HIV is able to cause indirect damage to neurons, with metabolic mechanisms playing a role in this process.

2.2.2 HIV impact on brain structure

Brain development is still affected by the persistent virus, even in the presence of cART. Basal ganglia calcification, atrophy within cerebral tissue and reductions in white matter (WM) and gray matter (GM) have been noted in HIV+ patients undergoing ART (Sarma et al., 2014; George et al., 2009). Damage to the myelin sheath within WM, together with the influx of supporting cells and a neuroinflammatory

response, is also observed following HIV infection (Chen et al., 2009; Anthony et al., 2005; Langford et al., 2002). Therefore, research is now being directed towards addressing the continued influence of HIV infection on cognitive health in individuals being treated for HIV, together with possible long-term harmful side effects of ART in the brain.

In children perinatally infected with HIV, a reduction in the corpus callosum volume, smaller left and right hippocampus structures and enlargement of the left and right nuclei accumbens have been reported previously, despite treatment (Yadav et al., 2017). Changes in the subcortical structure have also been found to associate with clinical measures of disease in adolescents following perinatal infection (Lewis-de los Angeles et al., 2016). Lewis-de los Angeles et al. (2016) reported an association between structural abnormalities in the thalamus in children with lower CD4 counts, while children with higher VL measures had abnormalities in the thalamus, globus pallidus, caudate and putamen structures.

2.2.3 Impacts of HIV exposure on brain development

A study carried out in Thailand and Cambodia, which utilized various tests to determine general and visual IQ scores, identified that children born to HIV+ women had lower visual IQs than unexposed children of a corresponding age (Kerr et al., 2014). Additionally, previous studies have revealed impeded development of important motor skills in HEU children aged 30-72 months (Van Rie, Mupuala & Dow, 2008). These findings imply that either exposure to antiretrovirals or to viral products, without being directly infected with HIV, can also have a negative effect on brain development.

2.2.4 HIV effects on brain function

Relatively few studies have been carried out to look at the impact of HIV on functional processes in the brains of children undergoing ART. In an fMRI study, increases in HIV RNA levels in perinatally-infected adolescents were found to correlate with greater blood oxygenation level dependent (BOLD) signaling within the default mode network (DMN), while connectivity between the DMN and the visual/executive networks, was significantly reduced (Herting et al., 2015). This is believed to indicate changes to network organization as a result of HIV infection (Herting et al., 2015), and may play a part in a compensatory process, as was also suggested to occur in HIV+ adults (Chang et al., 2001).

The task-based study by Chang et al. (2001) showed that HIV also affects working memory in adults. The process of carrying out tasks was slower in HIV+ individuals and they showed a greater BOLD response in the frontal lobe and parietal regions involved with working memory (Chang et al., 2001).

However, in addition to signaling being observed in similar regions to those of control individuals, greater signaling was observed specifically in the surrounding regions for HIV+ adults. (Chang et al., 2001). This is indicative of the need of HIV+ individuals to utilise their neural reserve, such as other signaling networks, to devote greater attention to carrying out tasks (Chang et al., 2001).

A resting-state functional MRI study of the long-term effects of HIV and ART on functional connectivity in perinatally infected children at 7 years, was recently conducted (Toich et al., 2018). Using independent component analysis to compare intra-network connectivity in HIV+ and uninfected (HIV-) children, no significant difference was found (Toich et al., 2018). However, through seed-based correlation analysis of whole brain connectivity, HIV+ children were found to have greater connectivity between the left superior temporal and right postcentral gyrus and between the left superior frontal gyri and the right anterior cingulate (Toich et al., 2018). Lower connectivity, contrastingly, was observed between 5 regions and their corresponding seeds: the left inferior frontal and the left cingulate gyri; the right supramarginal and the right frontal gyri; the right/middle occipital gyri and the right cuneus; the left medial frontal gyrus/anterior cingulate and the left cingulate gyrus; the left medial frontal/cingulate gyri and the left paracentral lobule (Toich et al., 2018). This indicates that HIV infection affects the development of functional connections in children (Toich et al., 2018).

These findings of structural and functional changes related to HIV infection, highlight the importance of addressing infections early.

2.3 MR Spectroscopy

2.3.1 MR spectroscopy in HIV studies

For HIV+ patients on ART who do not progress to a state of dementia, the effects of disease, which are more subtle at this stage due to treatment-driven viral suppression, are more clearly seen in terms of altered metabolic activity in specific regions of the brain, rather than structural changes (Paul et al., 2008).

HIV disease progression can effectively be tracked using 1H-MRS, while improvement in neurological health following an intervention, in this case ART, can be shown in terms of restored metabolic activity (Wilkinson et al., 1997). The metabolites typically studied using MRS include markers of neuronal and membrane integrity, N-acetyl aspartate (NAA) and choline (Cho) respectively, and compounds crucial for energy metabolism, creatine (Cr) and phosphocreatine (PCr) (for review see Soares & Law, 2009). Glutamate (Glu), an important excitatory neurotransmitter of the nervous system, can be measured together with its synthetic end-product, glutamine (Gln). Myo-inositol (Ins), a marker of glial

proliferation, but not restricted to glial cells alone, is also of interest in MRS studies. Measures of total choline (tCho; glycerophosphocholine + phosphocholine), total creatine (tCr; creatine + phosphocreatine), total NAA (tNAA; NAA + N-acetyl-aspartyl-glutamate) and Glx (Glu + Gln) are also of interest.

Glu is the principal excitatory neurotransmitter despite the fact that very little Glu is able to cross the BBB (Daikhin & Yudkoff, 2000). Glu is generated within the CNS, with glucose providing the source of carbon and other amino acids contributing the nitrogen component (Daikhin & Yudkoff, 2000). This mechanism is advantageous as it draws necessary resources into the brain. The activity of Glu is carefully regulated through cyclic Glu-Gln metabolism, where different steps are restricted to specific compartments (Bak, Schousboe & Waagepetersen, 2006). The uptake of Glu from the synaptic cleft is carried out by astrocytes and is important for lowering the metabolic concentration, such that a concentration gradient is maintained and greater signaling is favoured (Daikhin & Yudkoff, 2000). The astrocytes convert Glu to Gln, a compound which no longer drives signaling and therefore is preferable for transportation into the neurons (Daikhin & Yudkoff, 2000). Glutaminase, within the mitochondria of the neurons, can then hydrolyse Gln to re-establish a Glu reserve for use in synaptic signaling (Bak, Schousboe & Waagepetersen, 2006; Daikhin & Yudkoff, 2000).

Similarly, NAA undergoes cyclic activity, with NAA and NAAG being generated within the neurons and broken down in the oligodendrocytes and astrocytes, respectively (Baslow, 2010). NAA and Ins have been identified as osmolytes, regulating the flow of water between cells and the extracellular fluid within the brain (Dai et al., 2016; Baslow, 2010).

The transport of Ins into cells is coupled with sodium transport and this plays an important role in responding to alterations in fluid levels and impacts signaling (Dai et al., 2016).

Although some Cr within the brain is supplied by dietary intake, it is largely synthesized within the CNS (Braissant et al., 2010; Tachikawa et al., 2004). This process involves glial cells, in which the Cr is initially synthesized from glycine and L-arginine, and neurons, in which the Cr is then utilized by mitochondria (Tachikawa et al., 2004). PCr provides an important energy source and is catabolized to release Cr and phosphate for adenosine triphosphate (ATP) generation (for review see Braissant et al., 2011; Gualano et al., 2010). Cyclical activity, enabling continued ATP generation, occurs with the addition of phosphate to Cr in the mitochondria (Steen et al., 2010).

Sources of Cho are similarly obtained from food intake, although Cho can also be generated within the body from phosphatidylethanolamine (for review see Zeisel & Da Costa, 2009). Cho is a precursor to many important compounds such as acetylcholine, a neurotransmitter, and phospholipids, which

make up the backbone of cell membranes (for review see Tayebati et al., 2015; Sanders & Zeisel, 2007). Cho plays a role in transporting signals across cell membranes and in their structural stability (for review see Sanders & Zeisel, 2007).

Metabolite concentrations can be quantified in relation to endogenous or exogenous reference compounds, or to other metabolites of interest, to provide valuable insight into abnormalities in metabolic activity. This can enable the identification of tissue-specific shifts in metabolic activity associated with pathogenesis or neurological disorders, therefore, providing a useful diagnostic tool (Howe & Opstad, 2003; Lin, Crawford & Barker, 2003).

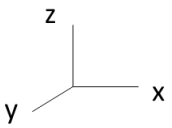
2.3.2 Basic physics underlying MRS

To experience nuclear magnetic resonance, nuclei require angular momentum, or non-zero spin. This occurs in nuclei containing an odd number of protons and/or neutrons. Due to this they have a non-zero magnetic dipole moment (MDM) which can be described in the form of a vector, with magnitude and direction, and experience torque in an external magnetic field (De Graaf, 2013).

On the application of a magnetic field, the MDMs tend to align with this field and precess around it (De Graaf, 2013). This can be thought of as the 'head' of the vector rotating in a circular motion, while the 'tail' is in a fixed position. The frequency of precession, ω_0, is described by the *Larmor* equation (equation 1), where γ is the gyromagnetic ratio constant and B0 is the strength of the magnetic field (tesla).

$$\omega 0 = \gamma B0 \quad (1)$$

The application of a radiofrequency (RF) pulse at the *Larmor* frequency of the precessing nuclei of interest (in this case protons) leads to the net magnetization M (the vector sum of MDMs) being shifted from its equilibrium position, aligned with the B0 field (z-axis), to a more horizontal position, where it precesses around B0 with a component in the xy plane. This occurs due to resonance and means that only small RF pulses need to be applied to generate an effect, as long as the frequency of the RF pulse is matched to the Larmor frequency (De Graaf, 2013). The larger the xy component of the M vector, the greater the signal that can be produced.



When the RF pulse is removed, the nuclei return to a low-energy state and energy is released, providing a signal that can be measured. The rate at which relaxation occurs can differ for different tissues, ultimately resulting in contrast between tissues. There are 2 types of relaxation that can occur. T1

relaxation corresponds with the recovery of the vertical (z) component of the M vector, while T2 describes the reduction of the horizontal (xy) component due to loss of phase coherence (De Graaf, 2013).

Valuable information regarding the environment of the nuclei and the structures in which nuclei exist, can be determined based on relaxation rates (De Graaf, 2013).

Chemical shifts result from the shielding effect caused by orbiting electrons, which varies according to the position of a nucleus within a molecule (De Graaf, 2013; Proctor & Yu, 1950). The effective magnetic field (B) experienced by the nucleus following shielding, can be described by equation 2, where σ is the shielding effect.

$$\mathrm{B} = \mathrm{B0}\,(1-\sigma) \tag{2}$$

The frequency of precession is proportional to the magnetic field amplitude and, therefore, changes due to shielding (De Graaf, 2013). It can now be described by equation 3.

$$\nu = (\gamma/2\pi)\,\mathrm{B0}\,(1-\sigma) \tag{3}$$

This allows certain metabolites within the brain to be identified according to their chemical shift (Govindaraju, Young & Maudsley, 2000), and metabolite concentrations can thus be quantified through MRS.

In order to select the regions for which MR spectra are obtained, there are two commonly used methods, Stimulated Echo Acquisition Mode (STEAM) and Point Resolved Spectroscopy (PRESS). STEAM and PRESS similarly utilize 3 RF excitation pulses followed by 'crusher' gradients to eliminate unwanted signals (De Graaf, 2013). A spectrum is obtained only for the region which experiences all 3 RF pulses.

In the case of STEAM, three 90° pulses allow very clean selection of voxels, but lead to high signal loss (De Graaf, 2013; Yongbi et al., 1995). PRESS, however, uses a sequence of 90°, 180°, 180°, referred to as a 'double spin-echo' approach (De Graaf, 2013). The 180° pulses refocus the signal and, thus, a better signal-to-noise ratio is achieved (De Graaf, 2013; Yongbi et al., 1995).

Due to the presence of large amounts of water throughout the body, signals from water tend to mask the signals from metabolites of interest from being detected by 1H-MRS. A chemical shift specific (CHESS) approach can thus be carried out, using a 90° RF pulse at a specific frequency to suppress the water signal such that spectra can be obtained for the metabolites of interest (Reiser, Semmler & Hricak, 2007; Haase et al., 1985).

2.3.3 Voxels of interest in HIV infection of the brain

A region of the brain which has been identified as a major target of HIV, is the basal ganglia (BG) (Berger & Arendt, 2000). Various roles for the BG have been identified, including interpreting language (Booth et al., 2007), in memory and in motor functions required for movement and learning (Wichmann & DeLong, 1996). The BG has been suggested to be involved in five distinct closed-loop signaling pathways within the subcortex of the frontal lobe, of which three play a role in behavior and executive functions - through the processing of various types of signals to enable a suitable response to be mounted (Bonelli & Cummings, 2007; Alexander, DeLong & Strick, 1986). The remaining two pathways have roles in motor activity (Alexander, DeLong & Strick, 1986).

The midfrontal gray matter (MFGM) is of interest as the frontal lobe is the last part of the brain to develop (Casey, Jones & Hare, 2008; Sowell et al., 1999) and carries out important cognitive and executive roles (Diamond, 2002). The metabolic activity in the frontal GM of HIV+ adults has been shown, through MRS, to be altered (López-Villegas, Lenkinski & Frank, 1997). A reduction in the ratio of NAA to Cr suggests that neuronal damage occurs here during HIV infection (López-Villegas, Lenkinski & Frank, 1997).

Peritrigonal white matter (PWM) is a region located posterior to the lateral ventricles, where a strong signal is achieved due to late myelination (Ackermann et al., 2014; Parazzini et al., 2002). This can be best detected in T2-weighted images (Parazzini et al., 2002). PWM is, therefore, a helpful region to analyse during neurological studies, as a clear distinction can thus be made between disease patients and normal control individuals (Liauw et al., 2008). As a region of WM in which radiologists had noted abnormalities during preliminary clinical studies of the children from the CHER trial (Ackermann et al., 2014), this is another region which may provide insight into the metabolic alterations which result in HIV+ children.

The BG, MFGM and PWM are, therefore, regions of interest in the brain when studying the effects of HIV and ART on neurodevelopment.

2.3.4 Previous MRS findings in HIV-positive patients

An initial study comparing metabolic activity in HIV+ adults before and following ART, revealed that HIV causes a reduction in the ratio of NAA to the sum of the metabolites of interest (NAA/NAA+Cho+Cr), within the WM of the parieto-occipital region of the brain (Wilkinson et al., 1997). Although an increase in this ratio was observed with treatment, full restoration was not achieved,

showing that metabolic activity is still hampered by HIV, in the presence of ARV drugs (Wilkinson et al., 1997).

2.3.4.1 Basal ganglia

An MRS study of HIV+ children with and without encephalopathy, found that children with HIV encephalopathy (HIVE) had lower NAA/Cr in the BG than children without encephalopathy (Pavlakis et al., 1995).

A higher viral load was found to be associated with lower concentrations of Cho, Cr and Ins in the BG of children (Keller et al., 2004). Further, a study looking at the ratio of NAA/Cho in the BG and subcortical WM, revealed a correlation between a better verbal IQ score for comprehension and arithmetic, and higher NAA/Cho concentrations (Gabis et al., 2006).

2.3.4.2 Frontal gray matter

Studies carried out in youth found increased Scy/Cr and Glu/Cr ratios in HIV+ subjects (Nagarajan et al., 2012) and only a trend in increased Cho/Cr was found during a chemical shift imaging study of the total GM within a select slice of the brain (Van Dalen et al., 2016).

2.3.4.3 White matter

Various studies have found elevated Cho/Cr in WM brain regions of HIV+ children in comparison to uninfected controls (Van Dalen et al., 2016; Prado et al., 2011), which Van Dalen suggested may indicate proliferation of glial cells. In a study by Keller et al. (2004), however, levels of Cho in the frontal WM were found to be lower for HIV+ children. This study also showed that HIV+ children had significantly different changes in NAA levels with age, in comparison to controls, which may provide evidence of the negative influence that HIV has on neuronal development in children (Keller et al., 2004).

2.3.4.4 Clinical measures and neurometabolic activity

$CD4^+$ T cell counts are generally considered the standard measure of disease progression and severity in individuals infected with HIV, however, CD8 counts also provide a useful measure and viral loads give a direct measure of the quantity of viral particles present within the host (Mulder et al., 1994; Mahalingam et al., 1993).

Previous studies have sought to investigate the relationship between neurometabolic activity and these various clinical measures. Lopez-Villegas et al. (1997) found that adults with a higher CD4 count, had significantly higher levels of NAA and Ins within cortical WM. A study by Lentz et al. (2009) found that reductions in the frontal cortex NAA levels were associated with an increase in CD8 T cell counts during the initial stages of HIV infection in adults. Although it cannot be concluded that a causal relationship exists between clinical measures and neurometabolic activity, it has been suggested, in both primate (Simian Immunodeficiency Virus) and human (HIV) studies, that changes in the peripheral immune response and viral load can impact, and/or be an indication of, the progression of disease within the brain (Lentz et al., 2009; Williams et al., 2005). Thus, changes in neurometabolic activity would be expected to associate with changes occurring exteriorly.

It is important to study the metabolic changes in children and more so to repeat studies throughout their years of development, as it has been established that metabolite concentrations differ in children with age (Holmes et al., 2017). It has been found that in some cases metabolite concentrations change in specific regions of the brain from childhood through to adulthood. For example, in a study by Pouwels et al. (1999) parietal gray matter NAA was found to increase and taurine decreased with age, whereas in the BG these metabolites remained unchanged with age. In the parietooccipital white matter, an age-related reduction in Gln and an increase in tNAA were detected (Pouwels et al., 1999).

In another study in healthy individuals, the ratio of Glu/Cr in the occipital and frontal cortices were found to decrease with age, from childhood into early adulthood (Shimizu et al., 2017). Therefore, if HIV alters the age-related metabolite changes that one would normally expect to observe in children, we would presume that differences in metabolic levels would exist between HIV+ and HIV- children in late childhood and adolescence. Thus, there are strong grounds for longitudinally assessing metabolic activity in HIV+ children as they enter adolescence.

2.4 Treating HIV in children

2.4.1 Initiation of ART in children

Although the earliest possible treatment of HIV+ children is generally favoured, toxic effects of ARV drugs on the brain (Robertson, Liner & Meeker, 2012; Cysique & Brew, 2009), and the evolution of viral populations to develop drug resistance under the selective pressure of ARV drugs (Richman, 1992), raise other important considerations when determining the optimal time for commencing treatment in infants.

The CHER trial was initiated to assess the most effective time for beginning cART in children in South Africa to establish control over viral infection and reduce infant mortality and, further to determine

treatment duration (Cotton et al., 2013; Violari et al., 2008). Although current protocols for treatment advise that children begin immediate treatment if they are found to be seropositive, this trial was initiated prior to the change in ART guidelines. Time-limited treatment was considered due to the limited supply of ARV drug regimens and medical resources in some rural areas of South Africa, where HIV+ children were needing treatment (Violari et al., 2008).

Children who tested positive for HIV were randomized into 3 groups under different treatment plans: deferred ART started only when clinical conditions were met (CD4% <25% or CDC stage C disease) , early ART (before 12 weeks of age) for 40 weeks (ART-40w) or early ART for 96 weeks (ART-96w) (Cotton et al., 2013; Violari et al., 2008).

A correlation was observed between the early initiation of ART and a reduction in infant mortality (Violari et al., 2008). The children who initiated early treatment also displayed better clinical outcomes (Violari et al., 2008). These findings challenged the WHO treatment recommendations of 2006, which had recommended that infants 11 months old or younger only initiate treatment when they reached CD4 percentages below 25 percent, or when they displayed serious clinical symptoms (WHO, 2006).

The CHER trial revealed that with early ART, fewer children needed to progress to the second-line regimen of ARV drugs (Cotton et al., 2013; Violari et al., 2008). The ART-96w children appeared to perform moderately better than the ART-40w children in terms of fewer children progressing to a second-line regimen and fewer clinical events being recorded for these children (Cotton et al., 2013). It has been suggested that longer treatment interruption periods would better enable the identification of patients in whom a functional or sterilizing cure may have been achieved (Cotton et al., 2013).

In another trial assessing when to begin treatment in children, the Pediatric Randomised Early versus Deferred Initiation in Cambodia and Thailand (PREDICT) trial, no differences in cognitive scores were seen in children over the age of one year who delayed treatment until their CD4 count dropped to less than 15%, compared to children who began ART with CD4 percentages between 15-30% (Puthanakit et al., 2013). These children did, however, differ when compared to HIV- (HEU and HU) children of a similar age (Puthanakit et al., 2013). In a continuation of this study of children from the PREDICT trial, structural differences between HIV+ and HIV- children were more evident in children below the age of 12 years (Paul et al., 2018).

A similar study by Cohen et al. (2016) showed abnormalities in structure and tissue architecture in the brains of HIV+ children 8-18 years old compared to uninfected controls (Cohen et al., 2016). A comparison of brain structure and tissue layout of HEU and HU children between the ages of 5-15 years

has shown no clear differences in neurodevelopment between these groups (Jahanshad et al., 2015). These studies indicate that the neurological damage inflicted by HIV occurs in the early stages of neurodevelopment, predominantly within the first year of life, and therefore suggest that early treatment is preferable.

2.4.2 The effect of early ART on neurocognitive health

Further study branching from the initial CHER trial looked more specifically at the influence of early ART on neurodevelopment in infants (Laughton et al., 2012). The children were assessed by clinicians based on the Griffiths Mental Development scales (GMDS), which had been approved for use among the South African population (Luiz, Foxcroft & Stewart, 2001). The clinicians were blinded to the treatment plan and HIV status of children, so as to minimise bias (Laughton et al., 2012). Better performance was observed in children on early ART than in those for whom treatment was deferred (Laughton et al., 2012). Although the findings of this study ascertained the importance of providing HIV+ infants with early treatment, as this favoured better short-term neurodevelopment, the long-term effects of this treatment strategy in the brain remained unclear.

To further investigate the long-term neurological effects of HIV infection, HIV+ children from the CHER trial (Cotton et al., 2013; Violari et al., 2008) and HIV-uninfected (exposed and unexposed) children from the same community, but involved in a Pneumococcal conjugate vaccine trial (Madhi et al., 2010), were recruited at the age of 5 into a longitudinal neuroimaging and neurocognitive study (Randall et al., 2017; Mbugua et al., 2016). HIV-exposed children had been involved in a prevention of MTCT (PMTCT) program, in which mothers were treated with a zidovudine-based ART regimen during pregnancy, in an attempt to reduce the risk of vertical HIV transmission (Violari et al., 2008).

At five years HIV+ children were found to no longer differ from HU children in terms of GMDS scores (Laughton et al., 2018). However, the cognition-related visual perception of HIV+ children was found to be significantly lower than controls (Laughton et al., 2018), suggesting that the ability of these children to interpret optic signals was, thus, impacted by the virus.

2.5 Longitudinal MRS in children from the CHER trial

A longitudinal MRS study, contributing to the larger neuroimaging study of children from the CHER trial, was initiated to assess how HIV infection impacted brain development in children undergoing cART, observing this at a metabolic level.

The findings of the MRS research carried out by Mbugua et al. (2016) between children from the CHER trial and control children at the age of 5 years, specifically assessing the BG region of the brain, revealed that lower levels of NAA and Cho occurred in uninfected children (all but 3 were HEU), than were observed in HIV+ children on early ART (Mbugua et al., 2016). The children who initiated ART within the first 3 months, were found to have higher Cho and NAA levels than children who initiated ART later (Mbugua et al., 2016) (Table 2.1). Higher NAA levels were interpreted as indicating less neuronal damage with early treatment. Upon examining the children for neurological disorders, a correlation was observed between higher Cho levels and neurological abnormalities (Mbugua et al., 2016). This suggests that early treatment is not completely preferential to late treatment for preventing damage caused by HIV.

Table 2.1: A table summarizing the findings of the analysis of absolute metabolite concentrations in the basal ganglia and midfrontal gray matter at 5, 7 and 9 years, as part of the longitudinal MRS study.

Voxel of interest	**Findings at 5 years (comparing early and late treatment) (Mbugua et al., 2016)**	**Findings at 7 years (Comparing HIV-positive and HIV-exposed-uninfected (HEU) groups to HIV-unexposed (HU) (Robertson et al., 2018)**	**Findings at 9 years (Comparing HIV-positive and HIV-exposed-uninfected (HEU) groups to HIV-unexposed (HU) (Robertson et al., 2018)**
Basal ganglia	↓ creatine, choline & glutamate in late vs early treatment group. ↑ NAA & choline in early treatment group, relative to HU children.	No significant difference across the HIV status groups.	↓ NAA and glutamate in HIV-positive and HEU groups. ↓ creatine and choline in HEU group.
Midfrontal gray matter (unpublished)	-	↓ glutamate in HIV positive children. ↑ choline in HIV positive and HEU groups.	↓ glutamate in the HIV positive group. Trend in ↓ choline in the HEU group.

MRS analysis in 7-year-old children surprisingly revealed no clear differences in basal ganglia metabolite levels of HIV+ children or HEU children, in comparison to the HU control children (Robertson et al., 2018) (Table 2.1). However, at 9 years, NAA and Glu levels were higher in HU children relative to the HIV+ and HEU children (Robertson et al., 2018) (Table 2.1). This finding suggests that possible damage to neurons, and reduced neuronal signaling occur in children exposed to or infected with HIV (Robertson et al., 2018). This result is similar to the reduced NAA that was previously observed in the BG of individuals with HIV encephalopathy (Pavlakis et al., 1995). At the age of 9 years, HEU children were found to have lower Cr and Cho levels than HU children (Robertson et al., 2018).

Robertson et al. (2018) have suggested that the reason for differing metabolic activity at 7 years and 9 years, may be due to the use of different scanners for the two age groups. Unlike the Allegra scanner (Siemens, Erlangen, Germany) with a single head coil, used to scan children at 7 years, at 9 years the children were scanned with the Skyra scanner (Siemens, Erlangen, Germany) which has 32 head coils, thus providing scans of greater quality (Robertson et al., 2018).

Lower CD4/CD8 count at enrolment was associated with reductions in Cho concentrations (5 and 7 years) or Ins levels (9 years) in the BG. Elevated Glu levels and lower Cr in the MFGM at 7 years were associated with increased CD4 percentages at enrolment, while at 9 years a reduction in Glu in the MFGM was associated with increased CD4 percentages at enrolment (Table 2.2).

Table 2.2: A summary of analysis of the relationship between absolute metabolite concentrations in the basal ganglia or midfrontal gray matter, and clinical measures at enrolment or at the time of scanning. Findings at the age of 5, 7 and 9 years are shown here.

Voxel of interest	**Findings at 5 years (Mbugua et al., 2016)**	**Findings at 7 years (Robertson et al., 2018)**	**Findings at 9 years (Robertson et al., 2018)**
Basal ganglia	↑CD8%, ↓CD4/CD8 count & ↓ CD4 at enrolment with ↓ NAA & ↓ choline. ↑ CD8 count at enrolment with ↓ NAA.	↓ CD4/CD8 at enrolment associated with ↓ choline levels	↓ CD4/CD8 at enrolment associated with ↓ myo-inositol.
Midfrontal gray matter (unpublished)	-	↑ CD4% at enrolment, CD4/CD8, age at ART initiation associated with ↑ glutamate. ↑ CD4% at enrolment associated with ↓ creatine.	↑ CD4% at enrolment associated with ↓ glutamate. ↑ CD4% at scan associated with ↓ choline.

Currently, the metabolic profiles of 5, 7 and 9-year-old children on early-ART have been characterized in different regions of the brain. However, more work could be carried out to assess neurometabolism longitudinally as these children grow older, particularly as certain regions of the brain continue to develop into adulthood. This would provide more insight into the long-term impact of ART and HIV persistence on brain development as they enter adolescence, a period where structural and hormonal changes are pronounced (Sisk & Zehr, 2005; Giedd, 2004).

2.6 Statistical analysis approaches

Linear regression analysis provides an advantageous method for comparing neurometabolic activity between children with different HIV statuses, and assessing associations between metabolic activity and clinical measures, as numerous variables can be controlled for in the models (Montgomery, Peck & Vining, 2012). Factors such as GM content within the selected voxel, age at scan and sex may contribute to differences observed between the groups and, thus, by adjusting for these variables, the effects of HIV status can specifically be identified apart from these other factors. However, this analysis only indicates how individual metabolite levels are impacted by HIV infection within the select regions.

There is growing interest regarding how neurometabolites interact with each other. Van Dalen et al. (2016) classified metabolites according to their primary roles, with NAA and Glu named 'neuronal' metabolites, and Ins and Cho labelled as 'glial' metabolites. They assessed the associations of these metabolites within GM and WM regions, as well as across these regions, through a linear regression approach (Van Dalen et al., 2016). They found that glial metabolites were associated in WM but not in GM, while neuronal metabolites associated with each other in both regions when controlling for age and HIV status (Van Dalen et al., 2016). A few significant relationships were also observed between glial metabolites and neuronal metabolites (Van Dalen et al., 2016). This approach, therefore, looked at metabolite patterns independently of the HIV status of individuals.

Other methods can be used to look at the relationships between metabolites and how these are altered by HIV infection. A study by Yiannoutsos et al. (2004), used factor analysis together with logistic regression to identify patterns of intra- and inter-regional metabolic activity that differentiated HIV+ individuals with ADC from HIV+ individuals without symptoms of dementia (Yiannoutsos et al., 2004). Factor analysis can be used to identify patterns between variables, based on their correlations. In this case, *varimax* rotation, an orthogonal rotation technique, was used (Yiannoutsos et al., 2004; Kaiser, 1958). This approach is advantageous particularly in exploratory studies, as it simplifies the interpretation of the findings (Abdi, 2003). However, it may not be accurate when there is a strong correlation between the factors generated. It may be preferable in exploratory factor analysis studies to additionally utilize an oblique rotation, such as *promax* rotation, as this allows for correlations to exist between the different factors. The findings from both approaches can then be reported and compared (Finch, 2006; Abdi, 2003). If there are strong correlations between factors, different results will be expected for these two approaches (Finch, 2006).

2.7 Aims & hypotheses of the study

2.7.1 Aims and objectives

The aim of this project is to cross-sectionally assess the long-term combined influence of HIV infection and early ART, or HIV exposure, on brain metabolism in children at 11 years, as part of a longitudinal study. To achieve this, steps will be taken to:

1. quantify various neurometabolites of interest in 11-year-old children, in terms of the absolute metabolite concentrations within specific regions of interest in the brain, and the ratio of these metabolites to total creatine.
   Using these concentrations:
   1.1 compare the metabolic profiles of HIV+ children to those of HU children
   1.2 compare the metabolic activity in HEU children to the HU controls.
2. assess the relationship between the various metabolites and clinical measures at enrollment and at the time of scanning.
3. investigate the metabolic patterns within the regions of interest, as well as across these regions, that characterize HIV+ children and distinguish them from HIV- children at 11 years.

The goal is to establish a better understanding of the underlying neurometabolic differences that occur at the age of 11 years in the brains of HIV+ children undergoing treatment, and in children exposed to HIV at the time of birth in comparison to controls. Further, we can consider what these differences could mean at a structural or functional level, while examining how these differences link to immune health. A novel component of this study is the use of factor analysis and logistic regression which has not previously been utilised in the study of spectroscopic data from children involved in the CHER trial.

2.7.2 Hypotheses

As there are few studies within an age range of 11-12 years, our hypotheses are dependent largely on findings in a cohort of these children at younger ages. Based on previous studies of the same cohort of children at the ages of 5, 7 and 9 years, we hypothesise similar trends in metabolic activity between the various HIV status groups to those previously seen. Lower levels of NAA and Glu may be expected in the BG of HIV+ children at the age of 11 years, than those observed in HU children of the same age. Cr and Cho levels are likely to be reduced in HEU children in relation to HU children, as observed at 9 years (Robertson et al., 2018). A reduction in the CD4/CD8 ratio at enrollment may be associated with a reduction of Ins or Cho in the BG at 11 years, as were seen at 9 and 7 years respectively (Robertson et al., 2018).

In the MFGM, HIV+ children at 11 years are expected to have lower Glu levels than HU children. Although higher Cho levels were observed at 7 years in HIV+ children, at 9 years Cho levels did not differ significantly from controls (unpublished data), and thus a similar trend of unaltered Cho levels would be expected at 11 years in HIV+ children relative to HU controls. Reduced Glu levels in the MFGM may correspond with higher CD4 percentages at enrollment (unpublished data).

3. Materials and methods

3.1 Participants

A total of 136 children from the same community were enrolled into this study, including 76 HIV+ children (11.0-12.5 years) participating in the Children with HIV Early Antiretroviral Therapy (CHER) trial (Violari et al., 2008), 30 HEU (11.0-12.1 years) and 30 HU children (11.3-12.1 years). Uninfected children had previously participated in a Pneumococcal conjugate vaccine trial (Madhi et al., 2010). HEU children were born to HIV+ mothers who had been involved in a PMTCT initiative and thus had undergone zidovudine treatment during pregnancy. These children also received zidovudine for a week and one dose of nevirapine following birth, with no further treatment (Robertson et al., 2018; Cotton et al., 2013).

A positive HIV diagnosis was confirmed by carrying out DNA and RNA PCR, for children over the age of 4 weeks (Cotton et al., 2013). HIV+ children were randomly assigned to 3 different treatment groups. Two groups were placed on early treatment (between ages 6-12 weeks) for a period of 40 weeks (ART-40w) or 96 weeks (ART-96w), while treatment was deferred in the third group (ART-def) (Cotton et al., 2013; Violari et al., 2008). The ART-def children began treatment once CD4 counts dropped below 20%, however this was then altered to 25%. A decision was made to re-evaluate all ART-def children, to determine whether they should initiate ART immediately, after the safety monitoring board for the CHER trial reviewed the data relating to infant mortality rates (Cotton et al., 2013). The board additionally suggested that no further children should be assigned to this arm (Cotton et al., 2013).

At the time of the 11-year scan, 56 HIV+ children were on a combination of Zidovudine, Lamivudine and Lopinavir/Ritonavir, while 7 children were on Abacavir, Lamivudine and Lopinavir/Ritonavir. There were 5 children taking a combination of Abacavir, Lamivudine and Efavirenz, 2 children receiving Zidovudine, Lamivudine and Efavirenz and 1 child was taking Zidovudine, Lamivudine and Abacavir. There were 2 children taking different combinations of 4 different drugs – Zidovudine or Abacavir with Darunavir, Ritonavir and Raltegravir, while treatment had been discontinued in 3 children.

3.2 Treatment strategy

The first line combination therapy regimen provided to HIV+ children consisted of protease inhibitors, lopinavir-ritonavir, and two nucleoside reverse transcriptase inhibitors, lamivudine and zidovudine (GlaxoSmithKline) (Cotton et al., 2013; Violari et al., 2008). All children in our sub-sample initiated treatment by 19 months of age. Viral load was suppressed in 61% of the HIV+ children by the age of 1 year and 74% by the age of 18 months (N=72). Children were placed on second line regimens, containing nevirapine/efavirenz, abacavir and didanosine in the event that an endpoint was reached,

such as children being diagnosed with severe CDC stage B or stage C disease, or if their CD4 counts dropped below 20% (Violari et al., 2008). When the children were scanned at 11 years, 97% of the children had suppressed viral loads.

3.3 Ethics approval

Ethics approval for the scanning of the children, and for use of the data in analysis, was obtained from the Human Research Ethics Committees at the University of Cape Town and Stellenbosch University. Parents provided written consent, while assent was obtained from the children.

3.4 Immune health assessment

Clinical measures of health, including CD4 counts, CD8 counts and viral loads, were taken at enrollment and at the time of scanning the children. CD4 percentage and the ratio of CD4 to CD8, provided further measures of immunological health. Viral load was categorized as high (>750,000 copies/mL), low (between 400 and 750,000 copies/mL) or suppressed (<400 copies/mL).

3.5 Neurological scanning

MRI scans were acquired at the Cape Universities Brain Imaging Centre (CUBIC), Groote Schuur Hospital, Cape Town. Scanning of the children was carried out at the age of 11.6 ± 0.3 years using the 3T Skyra scanner (Siemens, Erlangen, Germany), firstly, obtaining T1-weighted, structural images using a multiecho magnetization prepared rapid gradient echo (MEMPRAGE) sequence (van der Kouwe et al., 2008). Pulse sequence timings included inversion time (TI) 1100 ms, repetition time (TR) 2530 ms, echo times (TE= 1.69/3.54/5.39/7.24), a resolution of 1.0x1.0x1.0 mm^3 and 144 slices.

Voxels of interest (Figure 3.1), with areas of 1.5x1.5x1.5 cm^3, were selected for spectroscopic analysis. The BG, a region containing deep gray matter, predominantly included the caudate nucleus and internal capsule structures, as well as the putamen and globus pallidus (Robertson et al., 2018; Mbugua et al., 2016). This voxel consisted of an average of 58.47% of GM (Table 3.1). The MFGM voxel comprised of 85.92% of GM, while the PWM consisted of 79.19% WM (Table 3.1).

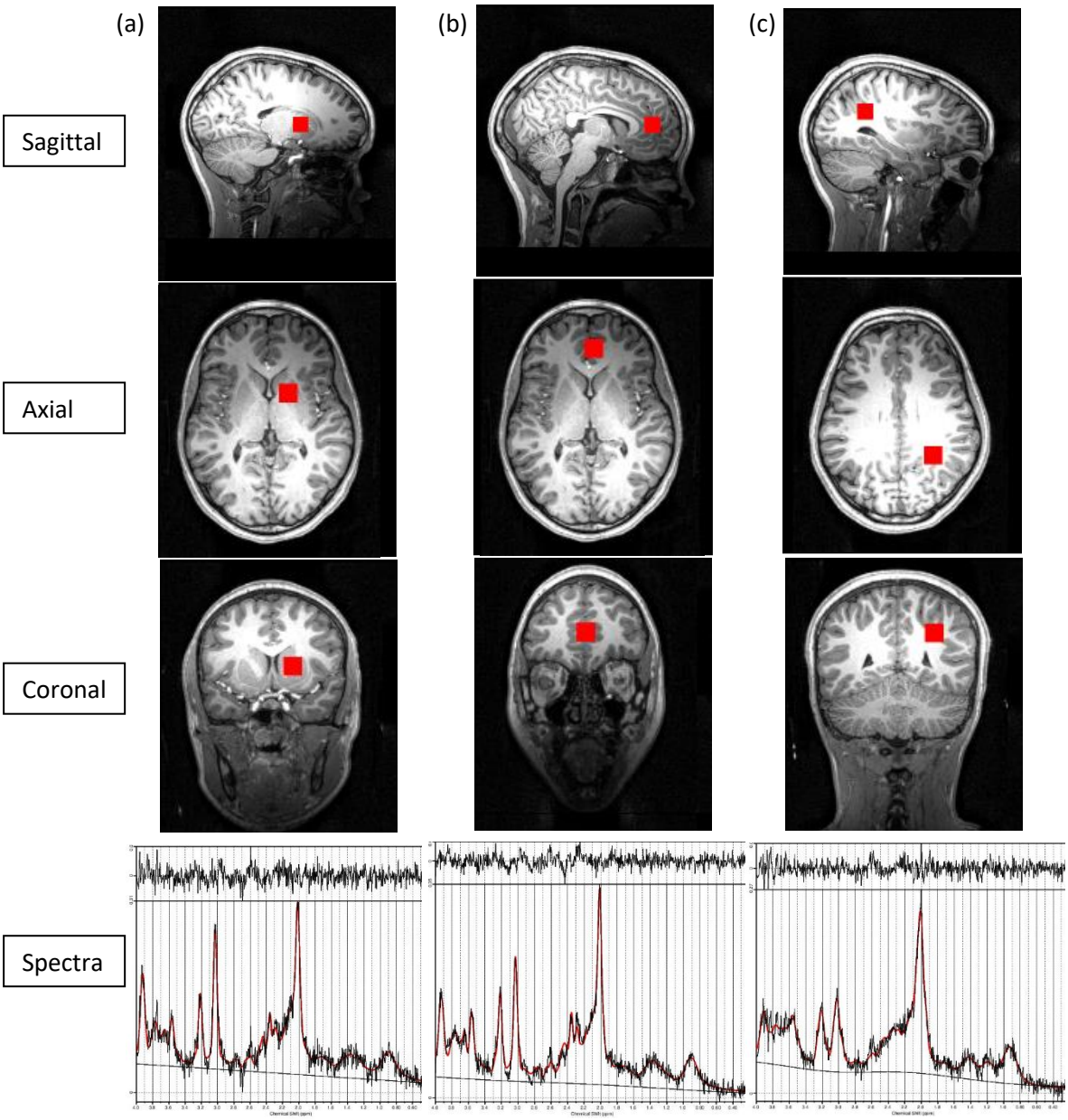


Figure 3.1: Structural MRI scans showing the positioning of the a) basal ganglia, b) midfrontal gray matter and c) peritrigonal matter voxels (1.5x1.5x1.5 cm^3) within the sagittal, axial and coronal planes, and their corresponding spectra.

Table 3.1: The mean gray matter (GM), white matter (WM) and cerebral spinal fluid content (CSF) (± standard deviation) of the regions investigated by 1H-MRS.

N=136	GM	WM	CSF
Basal ganglia	58.47±8.16%	41.39±8.19%	0.14±0.86%
Midfrontal gray matter	85.92±4.90%	2.45±1.47%	11.60±5.23%
Peritrigonal white matter	20.44±12.37%	79.19±12.42%	0.37±0.69%

Single voxel proton MR spectroscopy (1H-MRS) was performed, using a point resolved spectroscopy (PRESS) technique, to obtain metabolite spectra for selected regions within the BG, the MFGM and the PWM. A TE of 30 ms and a repetition time (TR) of 2000 ms were used, 64 averages, 1300Hz spectral bandwidth and vector size 1024 points, with chemical shift selective (CHESS) water suppression. A single additional scan was carried out without CHESS suppression, to obtain a water reference.

3.6 Processing the spectra

1H-MRS voxels were registered to the T1-weighted structural images. To determine the tissue composition and water concentration for each voxel, the structural image was segmented into WM, GM and cerebrospinal fluid (CSF) using Statistical Parametric Mapping (SPM12) software (http://www.fil.ion.ucl.ac.uk/spm).

The water concentrations of each voxel were calculated for each individual using the following equation:

$$\text{wconc} = (43300 * \text{GM fraction} + 35880 * \text{WM fraction} + 55556 * \text{CSF fraction}) / (1 - \text{CSF fraction})$$

This assumed a concentration of 55.6 mol/l in CSF, similar to that of pure water, a concentration of 43.3 mol/l in GM and a concentration of 35.9 mol/l in WM (Gasparovic et al., 2006).

Correction of eddy currents and baseline signal correction (Helms, 2008; Ernst, Kreis & Ross, 1993) were carried out using LCModel (version 6.3-1) as previously outlined (Provencher, 2001). A basis set of individual metabolite spectra, generated using the same scanning parameters, were used to fit the raw spectral data so that the metabolites could be quantified. The water-scaling method was utilized to determine the absolute concentrations of metabolites, as well as the ratios of metabolites to creatine (an intrinsic reference metabolite) (Helms, 2008; Ernst, Kreis & Ross, 1993).

The quality of spectral data was assessed in terms of the full width at half maximum (FWHM) and the signal-to-noise ratio (SNR) given by LCModel. Data with an FWHM less than 0.07 ppm or SNR greater than 7.0, which are considered to be of low quality, were excluded from analysis (as in Robertson et al., 2018). Scans were also excluded if the TE differed by more than 5ms from TE30.

Figure 3.2 provides a concise flow diagram summarizing the study participants, the data collected, spectral processing and analytical steps implemented in this study.

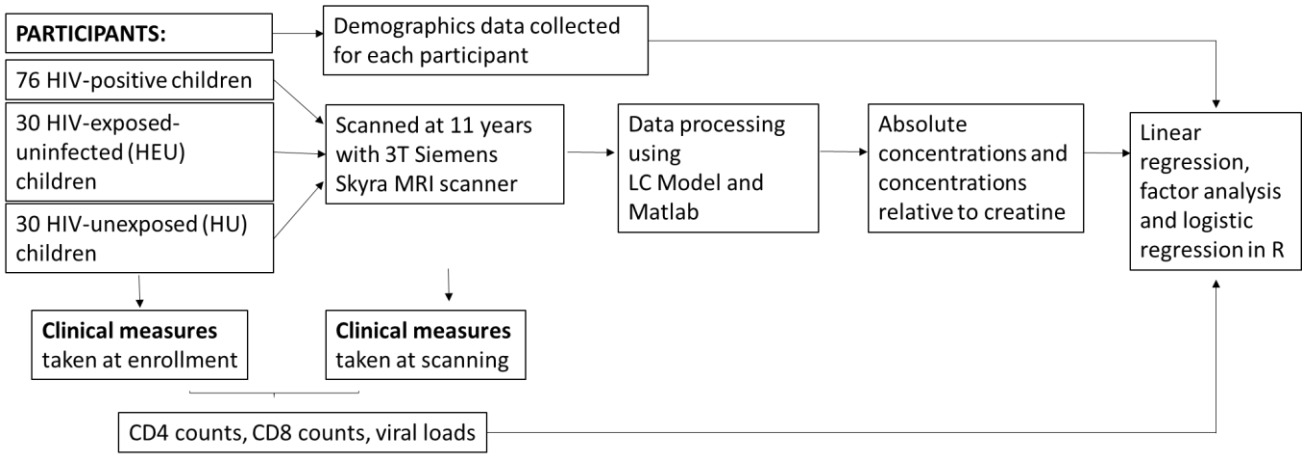


Figure 3.2: A pipeline summarising the steps involved in data collection, processing and statistical analysis carried out for this study.

3.7 Statistical analysis

3.7.1 Linear regression analysis

Statistical analysis was carried out using R software (R Core Team, 2018). Multiple linear regression models were used to examine HIV status group differences in absolute metabolite concentrations and the ratios of the metabolites to total creatine, controlling for potential confounding factors, such as age at time of scanning, sex, gray or white matter content and ethnicity (Equation given in Appendix A.1.1) (as in Robertson et al., 2018). These models were weighted using percentage standard deviation values obtained from LCModel. Ethnicity was determined based on the home language of children, isiXhosa or Afrikaans (for children of mixed ancestry), so as to account for culture and linguistics. Comparisons were made between HIV+ and HU children and between HEU and HU children. Further analysis was carried out to compare HIV+ and HIV- (HEU and HU) children.

The direct effect of sex was additionally investigated and no differences in metabolite concentrations were detected between males and females. However, sex was retained in the linear regression models due to the unbalanced number of males and females across the HIV status groups and to enable comparison to previous studies at younger ages, in which sex was controlled for.

3.7.2 Clinical measures and metabolite levels

In addition, the relationship between metabolite levels and immunological measures, such as the CD4 count and CD4/CD8 ratio at the time of enrolling children into the study, and at the time of scanning,

were assessed statistically through linear regression models - adjusting for age at scan, sex, ethnicity and GM/WM content. The relationship between metabolite concentrations and age at initiation of treatment, was also investigated. Clinical data at enrollment was defined as data which was collected before the age of 12 weeks, when the HIV+ children were enrolled into the CHER trial. Clinical data at scan was excluded from analysis if it was obtained more than 6 months from the date of scan.

3.7.3 Outliers and statistical significance

Outliers in metabolite and clinical data were identified as values which lay at 1.5 times the interquartile range (IQR), above or below the upper or lower quartiles respectively (Horn et al., 2001). Outliers were identified according to the entire cohort of children, rather than identifying outliers for each HIV status group individually. Analysis was carried out both including and excluding outliers. Outliers were excluded from the final models if the interpretation differed between these approaches.

P-values are reported to give an indication of the strength of findings for the readers benefit. However, statistical significance will not be assigned to results based on a p-value threshold of 0.05, as is traditionally done, as more statisticians are moving away from the use of this arbitrary value (Dahiru, 2008).

3.7.4 Pearson's correlation, factor analysis and logistic regression

For Pearson's correlation and factor analysis, metabolite concentrations were scaled and mean-centered to ensure that the values were on the same scale. Pearson's correlation analysis was carried out to assess crude relationships between the various metabolites both within the regions of interest, and across these regions. An r value greater than 0.5 was considered a strong correlation.

Exploratory factor analysis was initially carried out for 5 metabolites, tCho, tCr, tNAA, Glu and Ins, across the 3 regions of interest. Three metabolites, tCho, tNAA and tCr, were selected for further use in factor analysis to investigate the intra- and inter-regional interaction of these metabolites, due to their more robust measurements (low % standard deviation) and their association with HIV infection. Names were assigned to each factor based on the main contributing metabolites, as determined by the simple scoring approach, their roles and physiological indications.

Factor analysis was carried out in R software (R Core Team, 2018) using a maximum likelihood approach, to generate factors in which metabolite concentrations are grouped according to the strength of their correlations. Two different rotations, *varimax* (Yiannoutsos et al., 2004; Kaiser, 1958) and *promax* (Finch, 2006; Abdi, 2003), were carried out independently. The number of factors

generated was determined by the minimum number that could account for greater than 70% of the variability in the data.

Scores for each factor were then calculated based on the factor loadings, using two different approaches (Yiannoutsos et al., 2004):

Weighted approach: weighting each contributing metabolite within a factor according to its factor loading and dividing by the sum of factor loadings.

Simple approach: based on their factor loadings, allocating each metabolite solely to the factor to which it contributes the highest weight.

Factor scores for both scoring approaches, and both rotation techniques, were included as independent variables in logistic regression analysis models, with HIV status (HIV+ vs HIV-) as the outcome variable. The final model, adjusting for age at scan, sex, ethnicity and GM/WM content, included factors for which there was an association with HIV status (Equation given in Appendix A.1.2).

A total of three different logistic regression models were generated – the first two based on weighted factor scores for factor analysis with *varimax* rotation and *promax* rotation respectively. The third contained simple factor scores, which were the same in both *varimax* and *promax* rotation. Receiver operating characteristic (ROC) curves, defining the sensitivity and specificity of the models, were plotted for each of the logistic regression models. The area under the curve (AUC) percentages, measures of the capability of the models to predict the outcome (HIV status), were further determined.

Finally, linear regression was used to compare factor scores between the original 3 HIV status groups: between HIV+ and HU children and between HEU and HU children.

4. Results

This study contributes to a larger longitudinal study of the impact of HIV disease and ART on brain development in children from the CHER trial, however, the results presented here are specifically those for cross-sectional analysis of 1H-MRS at 11 years.

Firstly, the demographic and clinical data are introduced for the cohort of children in this study. Linear regression analysis of the individual metabolite levels within the BG, PWM and MFGM are then presented, followed by the linear regression models that investigate the relationship between the metabolites and various clinical measures. Raw metabolite correlations, not accounting for other factors, are then studied through Pearson's correlation analysis. Finally, factor analysis and logistic regression are used in conjunction to identify patterns of metabolic activity that are impacted in HIV+ individuals.

Table 4.1: Sample demographics.

N= 136	HIV positive	HEU	HU
N	76	30	30
Sex (Male)	36 (47%)	20 (67%)	15 (50%)
Age at scan (years)	11.62 (sd=0.27)	11.47 (sd=0.23)	11.67 (sd=0.24)
Ethnicity (Xhosa)	67 (88%)	29 (97%)	20 (67%)
Weight at scan (kg)	35.2 (sd=8.2)	38.1 (sd=11.8)	41.1 (sd=11.1)
Height at scan (cm)	140.0 (sd=9.0)	142.6 (sd=10.2)	145.5 (sd=8.9)

HIV-exposed-uninfected (HEU); HIV-unexposed (HU)

4.1 Participant demographics and clinical data

Table 4.1 provides a summary of the demographics for children included in this study. One-way analysis of variance (ANOVA) with post-hoc analysis revealed that HEU children were younger than HIV+ and HU children. Additionally, HIV+ children weighed less and were shorter than HU controls at the time of scanning.

As they did not meet the criteria described in the methods chapter, spectra for 3 participants were excluded from BG analyses (1 HIV+, 1 HEU and 1 HU), 4 from MFGM analyses (3 HIV+ and 1 HU) and 7 from analyses of PWM data (3 HIV+, 1 HEU and 3 HU). All spectra from these individuals, together with an HEU individual who consistently had influential results (identified from Cook's D statistic plots) for MFGM metabolite concentrations, were all excluded from Pearson's correlation analysis, factor analysis and logistic regression analysis.

In table 4.2, the clinical measures for HIV+ children are summarized. All HIV+ children began treatment before the age of 76 weeks. Treatment was interrupted in half of the children, from between the ages

of 45-115 weeks and for an average length of 80 weeks. A total of 40 children were classified as having CDC stage C of disease. At the time they were scanned for this study, 69 children had suppressed viral loads.

Table 4.2: A summary of clinical measures for HIV+ children at enrollment (before 12 weeks old) and at scan (11 years).

	N=76
At enrollment	
CD4 count (cells/mm^3)	1819 (sd=894)
CD4 %	32.9 (sd=10.5)
CD4/CD8 *	1.3 (sd=0.7)
Baseline viral load **	
High	41 (57%)
Low	31 (43%)
Suppressed	0
At scan	
CD4 count (cells/mm^3) ***	900 (sd=360)
CD4%***	37.8 (sd=6.8)
Viral load at scan****	
High	0
Low	2 (3%)
Suppressed	69 (97%)
cART initiation and interruption	
Age at starting cART (weeks)	14.7 (sd=13.4) range 5.9-75.7
Children with interrupted treatment	38 (50%)
Age at interruption (weeks)	
ART-40w (N=23)	49.6 (sd=2.7) range 45-56
ART-96w (N=15)	105.4 (sd=3.2) range 95-115
Duration of interruption (weeks)	80.4 (sd=103.0) range 5.7-398.0
CDC classification	
A	8 (11%)
B	14 (18%)
Severe B	14 (18%)
C	40 (53%)

* CD8 data at enrollment obtained after 12 weeks for 10 participants and missing for 3 participants.

** VL at enrollment missing for 4 children.

*** CD4 data for 8 participants was >6 months from scan and thus, excluded.

**** VL data at scan for 5 participants was >6 months from scan and thus, excluded.

4.2 BG metabolite concentrations in HIV+, HEU and HU children.

HIV+ children were found to have higher absolute tCho levels in the BG than HU control children at 11 years ($p = 0.03$; Table 4.3, Figure 4.1). Similar findings were observed in the ratios of the various metabolites to tCr, where an elevation of tCho/tCr was found in the BG of HIV+ children compared to the control group ($p = 0.02$; Table 4.4). There were no differences between HIV+ and HU children for other metabolites.

A comparison between the absolute metabolite concentrations in HEU and HU children, yielded no substantial differences in the BG region (Table 4.3). This too was confirmed by analysis of the ratios to tCr, for which no differences in BG metabolite ratios were found between HEU and HU children (Table 4.4).

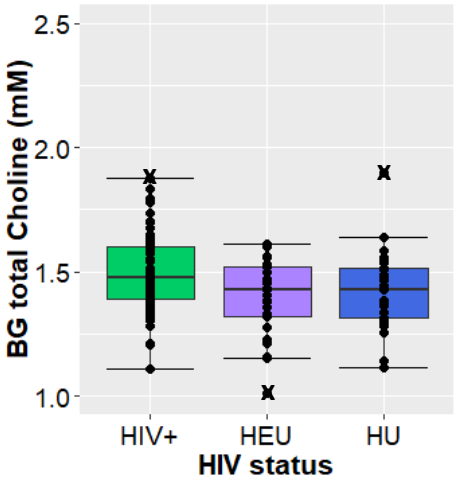


Figure 4.1: The absolute concentration of total choline in the basal ganglia shown as a dot and box plot (median and IQR) by diagnostic group at 11 years. (75 HIV+, 29 HEU, 29 HU). Outliers removed from regression analysis are indicated by an X.

Table 4.3: Linear regression analysis comparing absolute metabolite concentrations in the basal ganglia, midfrontal gray matter and peritrigonal white matter of HIV+ and HEU children, to HU children. Age at scanning, sex, ethnicity and gray/white matter content have been adjusted for.

	HIV+			**HEU**		
	B	SE	p	B	SE	p
Basal Ganglia (N=133: 75 HIV+, 29 HEU, 29 HU)						
tNAA	-0.12	0.16	0.47	-0.26	0.20	0.21
NAA	-0.23	0.15	0.13	-0.24	0.19	0.21
Glx	0.01	0.30	0.98	-0.12	0.37	0.74
Glu	0.13	0.21	0.53	0.15	0.26	0.58
tCho	0.07[a]	0.03	0.03	-0.01[a]	0.04	0.87
Ins	0.12	0.14	0.39	-0.19	0.17	0.28
tCr	-0.04	0.12	0.72	-0.08	0.15	0.62
Cr	0.06	0.20	0.77	-0.08	0.25	0.75
Midfrontal gray matter (N=132: 73 HIV+, 30 HEU, 29 HU)						
tNAA	0.09	0.16	0.55	-0.26[b]	0.20	0.20
NAA	0.06	0.16	0.70	-0.26[b]	0.20	0.20
Glx	-0.53[c]	0.30	0.08	-0.85[c]	0.38	0.03
Glu	-0.12	0.24	0.62	-0.44[b]	0.30	0.15
tCho	0.14	0.04	<0.001	0.00[b]	0.05	1.00
Ins	0.27	0.14	0.05	-0.26[b]	0.17	0.14
tCr	0.11	0.12	0.36	-0.15[b]	0.15	0.30
Cr	0.06	0.13	0.66	-0.34[b]	0.17	0.05
Peritrigonal white matter (N=129: 73 HIV+, 29 HEU, 27 HU)						
tNAA	-0.40	0.18	0.03	-0.54	0.23	0.02
NAA	-0.33[d]	0.16	0.05	-0.50	0.21	0.02
Glx	-0.44[e]	0.28	0.11	-0.18[e]	0.34	0.60
Glu	-0.28[f]	0.19	0.15	-0.22[f]	0.24	0.35
tCho	0.01	0.04	0.81	-0.05	0.05	0.33
Ins	-0.26	0.14	0.06	-0.26	0.17	0.13
tCr	-0.20	0.08	0.01	-0.17	0.10	0.09
Cr	-0.03	0.10	0.78	-0.05	0.12	0.68

Unstandardised coefficient (B); Standard error (SE) and p-value (p).
[a] 1 HIV+, 1 HEU and 1 HU outlier excluded from analysis; [b] 1 HEU participant excluded from all MFGM analysis due to being influential; [c] 1 HIV+, 1 HEU and 1 HU outlier excluded from analysis; [d] 1 HIV+ outlier excluded; [e] 1 HU outlier excluded; [f] 1 HIV+ and 1 HU outlier excluded.

Table 4.4: Results of linear regression analysis comparing the ratios of metabolites to total creatine (tCr) between unexposed children and HIV+ or HEU children, when adjusting for age at scan, sex, ethnicity and gray or white matter content.

	HIV+			**HEU**		
	B	SE	p	B	SE	p
Basal Ganglia						
tNAA/tCr	0.00	0.02	0.85	-0.02	0.03	0.41
NAA/tCr	-0.02	0.02	0.35	-0.02	0.02	0.41
Glx/tCr	0.01	0.03	0.80	0.00	0.04	1.00
Glu/tCr	0.03	0.03	0.29	0.03	0.03	0.32
tCho/tCr	0.01	0.00	0.02	0.00	0.01	0.46
Ins/tCr	0.02	0.02	0.38	-0.03	0.02	0.25
Midfrontal gray matter						
tNAA/tCr	0.00	0.02	0.99	-0.01	0.03	0.72
NAA/tCr	0.00	0.02	0.99	-0.01	0.02	0.76
Glx/tCr	-0.07	0.05	0.14	-0.05	0.06	0.38
Glu/tCr	-0.03	0.03	0.35	-0.02	0.04	0.67
tCho/tCr	0.02	0.01	0.001	0.01	0.01	0.42
Ins/tCr	0.03	0.02	0.09	-0.02	0.02	0.40
Peritrigonal white matter						
tNAA/tCr	0.00	0.04	0.92	-0.05	0.04	0.25
NAA/tCr	0.00	0.03	0.96	-0.04	0.04	0.35
Glx/tCr	-0.02	0.07	0.74	0.02	0.08	0.77
Glu/tCr	-0.01	0.04	0.78	0.00	0.05	0.98
tCho/tCr	0.02	0.01	0.06	0.00	0.01	0.86
Ins/tCr	-0.01	0.03	0.65	-0.02	0.04	0.56

Unstandardised coefficient (B); Standard error (SE) and p-value (p).

4.3 MFGM metabolite concentrations in HIV+, HEU and HU children.

In the MFGM region, tCho levels were higher in HIV+ children (p = <0.001; Table 4.3, Figure 4.2 a). This finding was complemented by an increase in tCho/tCr observed in HIV+ children compared to HU controls (p = 0.001; Table 4.4). Elevated Ins was additionally observed when looking at the absolute metabolite concentrations in HIV+ children compared to HU children (p = 0.05; Table 4.3, Figure 4.2 b), however only a trend for elevated Ins/tCr was seen when analysing the ratios of metabolites to creatine (p = 0.09; Table 4.4). HIV+ children had a trend for lower absolute Glx in the MFGM, however, analysis of the ratios to tCr yielded no such trend or difference compared to HU children. The levels of other metabolites were not noticeably different between HIV+ children and controls.

The levels of Glx in the MFGM also differed between HEU and HU children, with HEU children expressing lower Glx levels (p = 0.03; Table 4.3, Figure 4.2 c). A reduction in Cr was also observed in HEU children (p = 0.05; Table 4.3, Figure 4.2 d). However, no differences between these groups were evident in the ratio of Glx to tCr (Table 4.4).

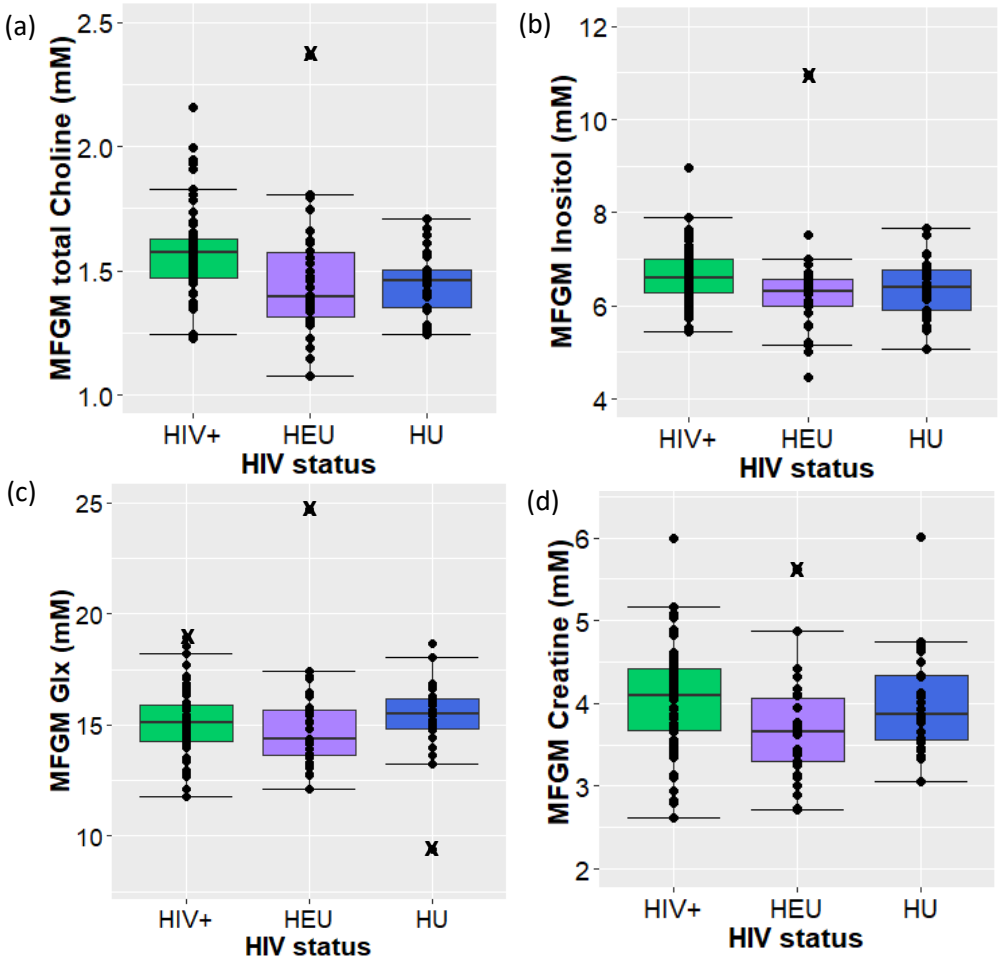


Figure 4.2: Dot and box plots (median and IQR) showing the absolute concentrations of (a) total choline, (b) inositol, (c) glutamate + glutamine (Glx) and (d) creatine in the midfrontal gray matter at 11 years, according to HIV status group (73 HIV+, 30 HEU, 29 HU). Outliers removed are indicated by an X.

4.4 PWM metabolite concentrations in HIV+, HEU and HU children.

In the PWM, HIV+ and HEU children were found to have a reduction in the concentration of tNAA (p = 0.03 & p = 0.02 respectively) and reduced NAA alone (p = 0.05 & p = 0.02 respectively), in comparison to HU controls (Table 4.3, Figures 4.3 a & b). Further, HIV+ children were found to have lower tCr levels

than HU children (p = 0.01), while only a trend for reduced levels of tCr (p = 0.09) was found between HEU and HU children (Table 4.3, Figure 4.3 c).

A trend for higher tCho/tCr was found for HIV+ children in the PWM (p = 0.06). The ratios of tNAA/tCr and NAA/tCr did not differ between HIV+ and HU controls, or between HEU children and controls (Table 4.4).

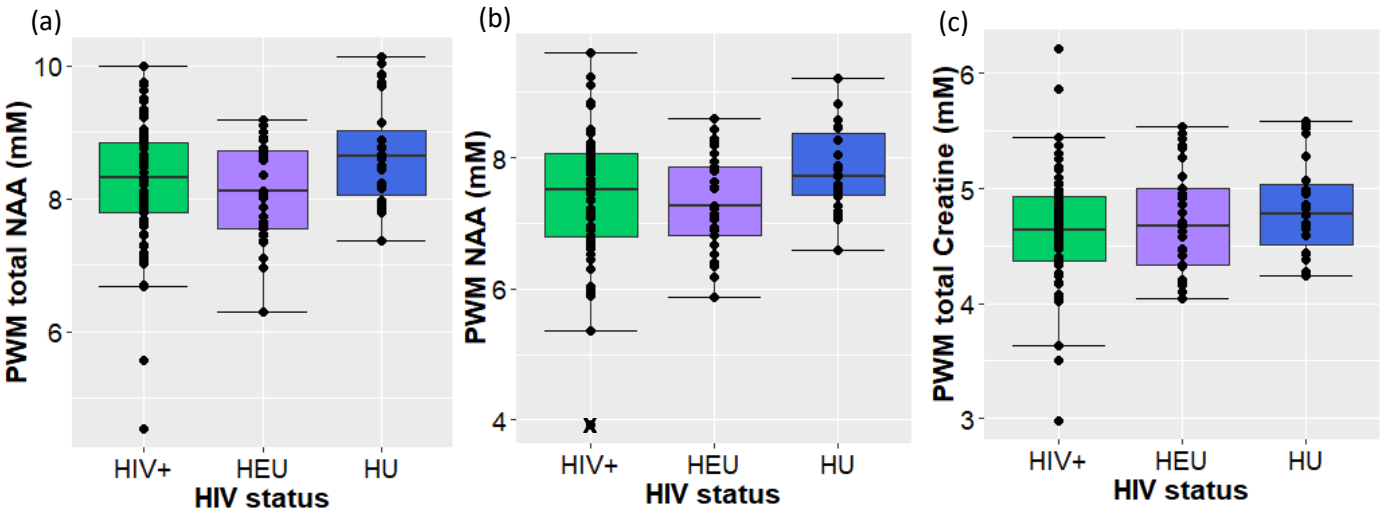


Figure 4.3: Box and dot plots (median and IQR) summarizing the absolute concentrations of (a) total NAA, (b) NAA and (c) total creatine in the PWM of children at 11 years, according to their HIV status (73 HIV+, 29 HEU, 27 HU). Outliers excluded from analysis are indicated by an X.

Elevated tCho in the MFGM of HIV+ children, was the only remaining finding of interest between HIV+ or HEU children and HU children after correcting for multiple comparisons.

4.5 Metabolite comparison between HIV+ and HIV- children

HIV+ children had higher absolute tCho in the BG (B=0.08; se=0.03; p=0.002) compared to the HIV- group, made up of all HEU and HU children combined (Supplementary table 1). A trend for elevated absolute Ins levels was also observed for HIV+ children (B=0.19; se=0.11; p=0.08).

Higher tCho (B=0.14; se=0.03; p <0.001) and elevated Ins (B=0.39; se=0.11; p <0.001) in the MFGM were found in HIV+ children compared to HIV- children (Supplementary table 1). Additionally, HIV+ children were found to have elevated absolute Cr levels (B=0.22; se=0.10; p=0.03). Lastly, no

metabolite levels were found to differ between HIV+ and HIV- children in the PWM (Supplementary table 1). However, a trend for lower tCr was observed in HIV+ children in this region.

4.6 Linear regression analysis for the relationship between clinical parameters and metabolite concentrations

Linear regression analysis assessing the relationship between metabolites in the 3 regions of interest and various clinical measures, led to a number of associations being found (Supplementary tables 2,3,4). In the BG, Ins was found to negatively associate with CD4 count at scan (B=-0.0007; se=0.0003; p=0.02), while Glu was positively associated with CD4/CD8 at enrollment (B=0.3137; se=0.1546; p=0.05) (Supplementary figure 1 a & b). In the MFGM, reductions in Ins (B=-0.0004; se=0.0002; p=0.03) and tCho (B=-0.0001; se=0.0001; p=0.01) were associated with higher CD4 count at scan and lower tCr was found to associate with CD4 count at enrollment (B=-0.0001; se=0.0001; p=0.01) (Supplementary figure 1 c, d & e). Finally, in the PWM, a negative relationship was observed between age at initiation of ART and NAA (B=-0.0022; se=0.0010; p=0.03) (Supplementary figure 1 f).

4.7 Intra- and inter-regional metabolite correlations

The correlations between metabolites within the same regions of interest and across these regions, were determined using Pearson's correlation analysis. Many strong intra-regional metabolite correlations were observed within the BG, MFGM and PWM, however fewer strong correlations were observed across the regions of interest for the entire cohort of children enrolled in the 11-year-old study (Table 4.5). The single inter-regional correlation which was found to be strong (r >0.5) across all the children was a positive correlation between BG Cho and MFGM Cho.

Table 4.5: Pearson's correlation coefficients (top-right) and p-values from correlation tests (bottom-left), showing the intra- and inter-regional metabolite correlations for the entire cohort of children (HIV+ and HIV-) at 11 years (N=123: 71 HIV+, 27 HEU, 25 HU).

	tCho BG	tNAA BG	tCr BG	Glx BG	Ins BG	tCho MFGM	tNAA MFGM	tCr MFGM	Glx MFGM	Ins MFGM	tCho PWM	tNAA PWM	tCr PWM	Glx PWM	Ins PWM
tCho BG		0.39	**0.62**	**0.55**	**0.55**	**0.52**	0.14	0.25	0.22	0.39	0.25	0.20	0.19	0.16	0.19
tNAA BG	<0.001		0.40	0.31	0.09	0.01	0.17	0.09	0.08	0.10	-0.07	0.36	0.19	0.11	-0.09
tCr BG	<0.001	<0.001		**0.68**	**0.55**	0.11	-0.04	0.17	0.09	0.17	-0.06	0.21	0.23	0.12	0.19
Glx BG	<0.001	<0.001	<0.001		0.44	0.07	0.00	0.09	0.22	0.18	-0.05	0.29	0.18	0.15	0.10
Ins BG	<0.001	0.31	<0.001	<0.001		0.35	0.02	0.27	0.13	0.47	-0.03	0.08	0.05	0.09	0.39
tCho MFGM	<0.001	0.96	0.22	0.42	<0.001		0.30	**0.54**	0.30	**0.63**	0.29	0.06	0.13	0.09	0.17
tNAA MFGM	0.12	0.07	0.65	0.98	0.83	0.001		**0.69**	**0.58**	0.43	0.20	0.26	0.27	0.12	0.13
tCr MFGM	0.01	0.34	0.06	0.33	0.002	<0.001	<0.001		**0.51**	**0.63**	0.24	0.18	0.35	0.10	0.23
Glx MFGM	0.01	0.41	0.35	0.02	0.15	0.001	<0.001	<0.001		0.47	0.14	0.25	0.31	0.31	0.20
Ins MFGM	<0.001	0.27	0.07	0.05	<0.001	<0.001	<0.001	<0.001	<0.001		0.22	0.19	0.19	0.10	0.41
tCho PWM	0.01	0.42	0.51	0.60	0.74	0.001	0.03	0.01	0.11	0.01		0.02	-0.10	-0.24	0.19
tNAA PWM	0.03	<0.001	0.02	0.001	0.41	0.55	0.004	0.04	0.005	0.04	0.86		**0.61**	0.48	0.04
tCr PWM	0.03	0.03	0.01	0.05	0.56	0.17	0.002	<0.001	<0.001	0.03	0.27	<0.001		**0.61**	0.30
Glx PWM	0.09	0.25	0.17	0.11	0.32	0.31	0.20	0.25	<0.001	0.27	0.01	<0.001	<0.001		0.13
Ins PWM	0.04	0.35	0.04	0.30	<0.001	0.06	0.17	0.011	0.03	<0.001	0.03	0.66	0.001	0.16	

Within-region correlations are indicated in blue; Across-region correlations >0.3 are shown in yellow and >0.5 in green.

A further exploration of the Pearson's correlations separately for HIV+ and HIV- individuals led to new, strong cross-regional correlations being identified.

The positive BG Cho/MFGM Cho correlation, as seen for the entire cohort, was particularly strong in HIV+ children (Table 4.6), while in HIV- children this correlation was weaker (r <0.3) (Table 4.7; Supplementary figure 2a).

Table 4.6: Results of Pearson's correlation analysis between metabolites within the voxels of interest and across these voxels, showing the correlation coefficients (top-right) and p-values (bottom-left) only for HIV+ children (N=71).

	tCho BG	tNAA BG	tCr BG	Glx BG	Ins BG	tCho MFGM	tNAA MFGM	tCr MFGM	Glx MFGM	Ins MFGM	tCho PWM	tNAA PWM	tCr PWM	Glx PWM	Ins PWM
tCho BG		0.34	**0.63**	**0.52**	**0.54**	**0.57**	0.08	0.24	0.16	0.35	0.18	0.14	0.12	0.12	0.12
tNAA BG	0.004		0.36	0.39	0.16	0.13	0.28	0.21	0.15	0.25	-0.11	0.28	0.17	-0.08	0.01
tCr BG	<0.001	0.002		**0.65**	**0.59**	0.13	-0.07	0.15	0.00	0.08	-0.14	0.07	-0.04	-0.14	0.10
Glx BG	<0.001	0.001	<0.001		0.40	-0.02	-0.09	0.00	0.11	0.06	-0.12	0.31	0.06	0.02	-0.02
Ins BG	<0.001	0.19	<0.001	0.001		0.34	-0.10	0.18	0.04	0.38	-0.20	0.10	-0.01	0.09	0.29
tCho MFGM	<0.001	0.30	0.27	0.90	0.004		0.31	**0.56**	0.36	**0.58**	0.25	0.01	0.09	0.14	0.22
tNAA MFGM	0.51	0.02	0.58	0.46	0.42	0.01		**0.74**	**0.63**	0.45	0.23	0.18	0.36	0.10	0.21
tCr MFGM	0.04	0.08	0.22	0.97	0.13	<0.001	<0.001		**0.52**	**0.62**	0.22	0.18	0.38	0.08	0.26
Glx MFGM	0.20	0.21	0.97	0.36	0.75	0.002	<0.001	<0.001		0.45	0.14	0.24	0.31	0.29	0.14
Ins MFGM	0.003	0.04	0.51	0.60	0.001	<0.001	<0.001	<0.001	<0.001		0.22	0.18	0.19	0.10	0.40
tCho PWM	0.14	0.38	0.24	0.32	0.10	0.04	0.05	0.06	0.26	0.07		0.04	-0.01	-0.22	0.22
tNAA PWM	0.26	0.02	0.57	0.01	0.43	0.91	0.14	0.14	0.05	0.14	0.76		**0.62**	**0.46**	0.02
tCr PWM	0.31	0.16	0.73	0.63	0.95	0.44	0.002	0.001	0.01	0.12	0.97	<0.001		**0.54**	0.31
Glx PWM	0.34	0.49	0.24	0.85	0.48	0.24	0.40	0.52	0.01	0.39	0.06	<0.001	<0.001		0.09
Ins PWM	0.30	0.94	0.40	0.90	0.01	0.06	0.08	0.03	0.26	0.001	0.07	0.89	0.01	0.48	

Within-region correlations are indicated in blue; Across-region correlations >0.3 are shown in yellow and >0.5 in green.

In HIV- children, strong positive correlations were found between Ins across the 3 regions (BG-MFGM, BG-PWM and MFGM-PWM) (Table 4.7). Additionally, the correlation between Cr in the BG and PWM was positive and strong in these children. In the HIV+ children, by contrast, the BG Cr/PWM Cr correlation was negative and weak (Table 4.6; Supplementary figure 2b). Further, the correlations between Ins across the regions of interest, although not entirely weak, were not strong as in HIV- children. A greater number of correlations with intermediate strength, were also observed in HIV- children compared to those seen in HIV+ children, or those seen overall.

Table 4.7: A matrix of Pearson's correlation coefficients (top-right) and the corresponding p-values from correlation tests (bottom-left), summarizing the interactions of metabolites within the regions of interest and across these regions in HIV-uninfected children (N=52).

	tCho BG	tNAA BG	tCr BG	Glx BG	Ins BG	tCho MFGM	tNAA MFGM	tCr MFGM	Glx MFGM	Ins MFGM	tCho PWM	tNAA PWM	tCr PWM	Glx PWM	Ins PWM
tCho BG		**0.50**	**0.66**	**0.58**	**0.54**	0.29	0.15	0.17	0.32	0.30	0.27	0.36	0.39	0.33	0.37
tNAA BG	<0.001		0.43	0.26	0.03	-0.11	0.07	-0.05	0.01	-0.02	-0.04	0.46	0.23	0.28	-0.18
tCr BG	<0.001	0.001		**0.74**	**0.54**	0.08	-0.03	0.22	0.17	0.26	0.02	0.38	**0.56**	0.42	0.29
Glx BG	<0.001	0.06	<0.001		**0.50**	0.13	0.10	0.23	0.38	0.31	0.03	0.28	0.41	0.37	0.29
Ins BG	<0.001	0.83	<0.001	<0.001		0.29	0.17	0.43	0.29	**0.59**	0.21	0.06	0.20	0.16	**0.61**
tCho MFGM	0.04	0.44	0.57	0.35	0.04		0.19	0.46	0.25	**0.60**	0.27	0.21	0.32	0.20	0.25
tNAA MFGM	0.30	0.62	0.86	0.49	0.22	0.18		**0.61**	**0.51**	0.35	0.11	0.40	0.17	0.19	0.05
tCr MFGM	0.22	0.73	0.12	0.10	0.002	0.001	<0.001		**0.53**	**0.62**	0.22	0.24	0.37	0.24	0.26
Glx MFGM	0.02	0.95	0.23	0.01	0.04	0.08	<0.001	<0.001		**0.54**	0.15	0.28	0.32	0.35	0.29
Ins MFGM	0.03	0.89	0.06	0.03	<0.001	<0.001	0.01	<0.001	<0.001		0.14	0.28	0.30	0.22	**0.55**
tCho PWM	0.05	0.77	0.91	0.86	0.13	0.05	0.44	0.12	0.28	0.33		0.01	-0.22	-0.23	0.21
tNAA PWM	0.01	0.001	0.01	0.04	0.66	0.14	0.003	0.09	0.05	0.05	0.96		**0.58**	0.50	0.06
tCr PWM	0.004	0.10	<0.001	0.002	0.15	0.02	0.24	0.01	0.02	0.03	0.11	<0.001		**0.72**	0.27
Glx PWM	0.02	0.04	0.002	0.01	0.27	0.15	0.17	0.08	0.01	0.11	0.10	<0.001	<0.001		0.15
Ins PWM	0.01	0.20	0.04	0.04	<0.001	0.07	0.71	0.06	0.04	<0.001	0.14	0.68	0.05	0.28	

Within-region correlations are indicated in blue; Across-region correlations >0.3 are shown in yellow and >0.5 in green.

4.8 Factor analysis and logistic regression

4.8.1 Factor analysis with *varimax* rotation

Factor analysis with *varimax* rotation, involving 5 metabolites from the 3 regions of interest, yielded 8 factors which accounted for >70% variability in the data (Supplementary table 5). Metabolites were grouped in a very region-specific manner here.

In further analysis using only 3 selected metabolites, tNAA, tCho and tCr, 5 factors, generated through factor analysis with *varimax* rotation, were found to account for 71.0% of variation within the data (Table 4.8). An inter-regional coupling of Cho across the BG, MFGM and PWM determined one factor (the multiregional inflammatory factor). Two factors consisted of separate, intra-regional couplings of NAA and Cr within the MFGM and PWM and were named the MFGM neuronal and PWM axonal factors, respectively. In white matter, NAA is a marker of axonal integrity. The fourth factor was driven

by tCr in the BG (BG energy factor), while the final factor was based on tNAA levels within the BG (BG neuronal factor).

Table 4.8: Factor loadings assigned according to correlations between total choline (tCho), total N-acetyl-aspartate (tNAA) and total creatine (tCr) within the basal ganglia (BG), midfrontal gray matter (MFGM) and peritrigonal white matter (PWM), when using *varimax* rotation.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
tCho BG	**0.78**	<0.1	0.15	0.48	0.37
tNAA BG	<0.1	<0.1	0.15	0.23	**0.64**
tCr BG	<0.1	<0.1	0.12	**0.95**	0.26
tCho MFGM	**0.68**	0.33	<0.1	<0.1	-0.11
tNAA MFGM	0.18	**0.71**	0.16	-0.14	0.22
tCr MFGM	0.30	**0.92**	0.18	0.15	<0.1
tCho PWM	**0.39**	0.17	-0.12	<0.1	<0.1
tNAA PWM	<0.1	0.12	**0.58**	<0.1	0.37
tCr PWM	<0.1	0.18	**0.98**	0.11	<0.1
	Multiregional inflammatory factor	MFGM neuronal factor	PWM axonal factor	BG energy factor	BG neuronal factor

The highest loading for each metabolite is shown in bold. Factors were named based on these metabolites as per the simple scoring approach.

The inclusion of weighted factor scores in logistic regression models led to 2 factors, the multiregional inflammatory factor and the PWM axonal factor, being found to significantly differentiate between HIV+ and HIV- individuals at 11 years. Therefore, only these factors were retained in the final model, adjusting for age at scan, GM/WM content, ethnicity and sex. The MFGM neuronal factor, BG energy factor and BG neuronal factor scores were excluded as they were not associated with the outcome, namely whether an individual had HIV or not.

In a logistic regression model including the weighted factor scores from factor analysis with *varimax* rotation, the odds of HIV infection were found to be 5.6 times greater for a unit increase in the multiregional inflammatory factor score, when adjusting for other confounding variables and the PWM axonal factor (Table 4.9). A unit increase in the PWM axonal factor was found to lead to a 65.5% reduction in the odds of HIV infection when adjusting for the multiregional inflammatory factor and other confounders.

Table 4.9: The results of a logistic regression model investigating the association between HIV status and the **weighted** inflammatory factor and peritrigonal white matter (PWM) axonal factor scores, based on *varimax* rotation, when adjusting for confounding variables (N=123: 71 HIV+, 52 HIV-).

	Odds ratio	95% CI	p-value
PWM axonal factor	0.35	0.14-0.80	0.02
Inflammatory factor	5.64	2.56-14.22	<0.001
Age at scan	2.26	0.41-12.19	0.33
BG gray matter content*	1.03	0.97-1.09	0.38
MFGM gray matter content*	0.96	0.86-1.08	0.52
PWM white matter content*	0.97	0.91-1.02	0.23
Sex (Male)	0.66	0.28-1.53	0.33
Ethnicity (Afrikaans)	0.37	0.11-1.19	0.10

* These refer to the percentages of gray/white matter that constitute these voxels, as determined by segmentation.

The introduction of simple factor scores showed a similar result, but a smaller effect, with a unit increase in the multiregional inflammatory factor score being associated with only 1.6 times higher odds of HIV infection (Supplementary table 6). A 31.4% decrease in the odds of HIV infection, was associated with a unit increase in the PWM axonal factor when adjusting for other confounding variables.

4.8.2 Factor analysis with *promax* rotation

The use of *promax* rotation in factor analysis, similarly led to a selection of 5 factors, which accounted for 70.8% variability in the data. Although factor loadings were not identical to those from the *varimax* rotation approach, these factors contained the same main contributing metabolites (Table 4.10).

Again, the multiregional inflammatory factor, together with the PWM axonal factor, were found to significantly associate with HIV status when carrying out logistic regression analysis (Table 4.11). In the case of weighted scoring, a single unit increase in the multiregional inflammatory factor was associated with 3.6 times higher odds of HIV infection when adjusting for the PWM axonal factor and other confounding variables. Adjusting for the multiregional inflammatory factor and other confounders, there was a 50.3% decrease in the relative odds of HIV infection for a unit increase in the PWM axonal factor.

The simple scoring approach yielded the same factor scores and logistic regression results for *varimax* and *promax* rotation techniques (see Tables 4.8 and 4.10).

Table 4.10: Factor loadings from factor analysis with *promax* rotation, for total choline (tCho), total N-acetyl-aspartate (tNAA) and total creatine (tCr) in the basal ganglia (BG), midfrontal gray matter (MFGM) and peritrigonal white matter (PWM).

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
tCho BG	**0.92**	-0.22	<0.1	0.18	0.29
tNAA BG	<0.1	<0.1	-0.14	0.11	**0.71**
tCr BG	<0.1	0.17	<0.1	**0.95**	0.13
tCho MFGM	**0.63**	0.27	<0.1	<0.1	-0.19
tNAA MFGM	<0.1	**0.69**	<0.1	-0.17	0.29
tCr MFGM	<0.1	**1.04**	<0.1	0.19	-0.12
tCho PWM	**0.36**	0.11	-0.16	-0.17	<0.1
tNAA PWM	<0.1	<0.1	**0.46**	<0.1	0.36
tCr PWM	<0.1	<0.1	**1.04**	<0.1	-0.12
	Multiregional inflammatory factor	MFGM neuronal factor	PWM axonal factor	BG energy factor	BG neuronal factor

The highest loading for each metabolite is indicated in bold.

Table 4.11: The results of a logistic regression model investigating the association between HIV status and the **weighted** inflammatory and peritrigonal white matter (PWM) axonal factor scores, based on *promax* rotation, when adjusting for confounding variables (N=123: 71 HIV+, 52 HIV-).

	Odds ratio	95% CI	p- value
PWM axonal factor	0.50	0.25-0.92	0.03
Inflammatory factor	3.62	1.95-7.43	<0.001
Age at scan	2.82	0.55-15.50	0.22
BG gray matter content*	1.02	0.96-1.08	0.54
MFGM gray matter content*	0.93	0.83-1.04	0.21
PWM white matter content*	0.96	0.90-1.02	0.20
Sex (Male)	0.64	0.27-1.47	0.29
Ethnicity (Afrikaans)	0.35	0.10-1.14	0.09

* These refer to the percentages of gray/white matter that constitute these voxels, as determined by segmentation.

Based on ROC curves for the logistic regression models including factor scores, the predictive abilities of the models were assessed. The logistic regression models including weighted scores, had AUC percentages of 77.4% for *varimax* rotation and 76.7% for *promax* rotation respectively (Figure 4.4 a & b). The model including simple scores had an AUC of 75.9% (Figure 4.4 c).

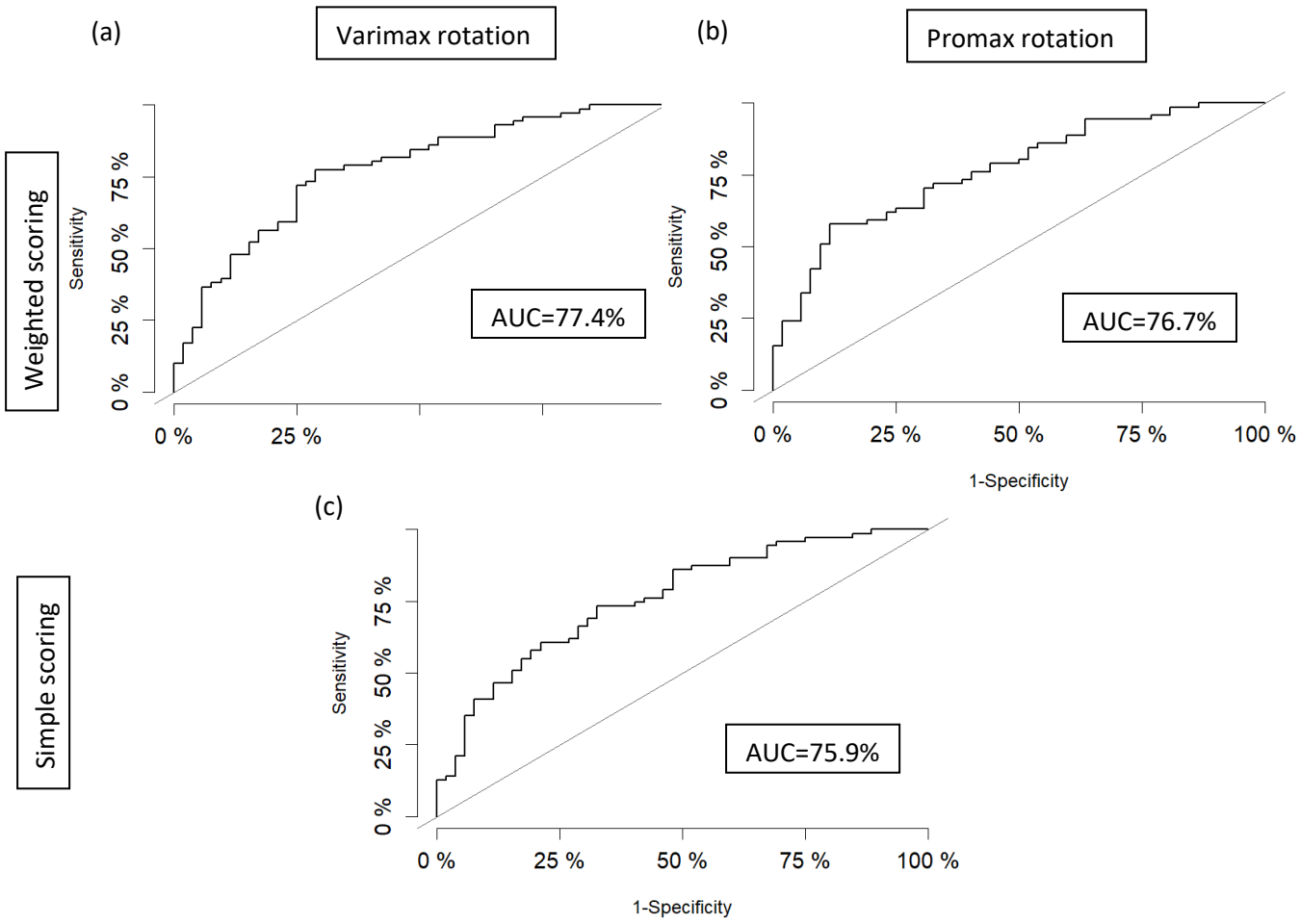


Figure 4.4: Receiver operating characteristic (ROC) curves showing the sensitivity and specificity of the logistic regression models created using factor scores based on factor analysis using (a) *varimax* or (b) *promax* rotation and generated by the weighted scoring approach or (c) the simple scoring approach: identical for *varimax* and *promax* rotations. Area under the curve (AUC) percentages are provided.

4.9 Comparison of factor scores between HEU and HU children

Additional linear regression analysis comparing simple factor scores between HEU and HU children, while adjusting for confounding variables, revealed that these groups did not differ in terms of the multiregional inflammatory score (B=-0.63; se=0.64; p=0.33) (Supplementary table 7 & Figure 4.5 a). However, HEU children had lower PWM axonal factor scores (B=-1.12; se=0.41; p=0.01) (Supplementary table 7 & Figure 4.5 b).

As expected, HIV+ children had elevated multiregional inflammatory (B=1.14; se=0.52; p=0.03) (Supplementary table 7 & Figure 4.5 a) and PWM axonal factor scores (B=-0.88; se=0.33; p=0.01) (Supplementary table 7 & Figure 4.5 b).

In both HIV+ and HEU children, no differences in MFGM neuronal factor, BG neuronal factor or BG energy factor scores were found in comparison to HU control children.

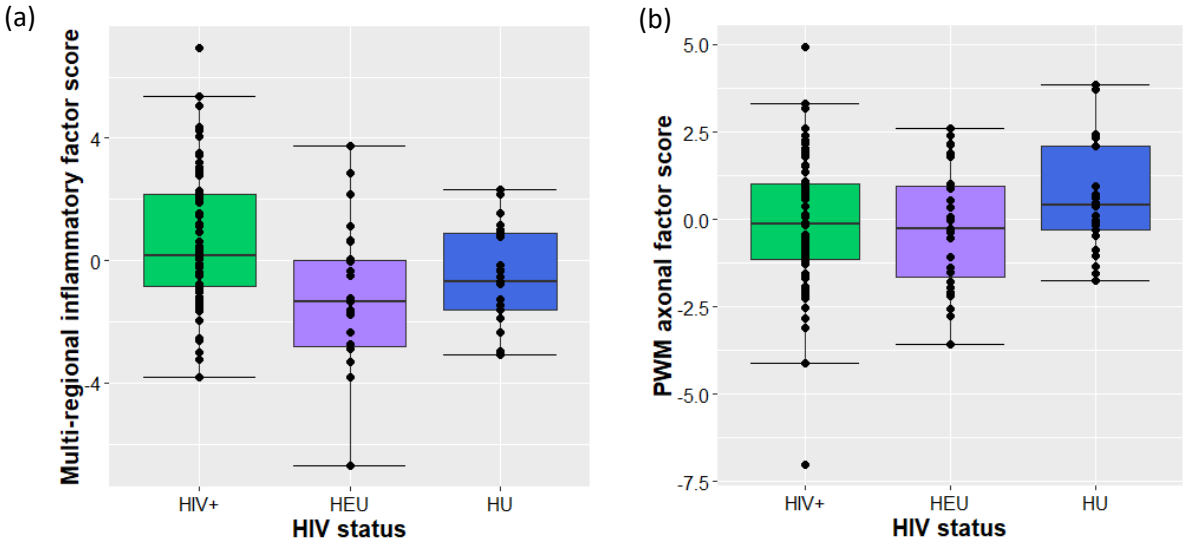


Figure 4.5: Simple factor scores according to HIV status groups, showing elevated multiregional inflammatory factor scores in HIV+ children and lower PWM axonal factor scores in HIV+ and HEU children compared to HU control children.

5. Discussion

This project sought to cross-sectionally assess the neurometabolic profiles of HIV+ and HEU children, following perinatal HIV infection and approximately 11 years of treatment or 11 years after exposure, respectively. In addition to examining the effects of HIV exposure, our goal was to address the neurological impacts of HIV in a cohort of South African children who were placed on the same ART regimen before the age of 19 months. The hope was to gain a better understanding of what chemical changes are occurring and to identify biomarkers of HIV infection in perinatally infected children at 11 years.

Focusing on biochemical activity in the BG, MFGM and PWM, we identified metabolic changes that occur in these specific regions of the brain compared to HU controls. Further, we investigated the relationships between the various metabolite concentrations and clinical indicators of HIV disease severity, both at enrollment and at time of scanning. Finally, we identified patterns of metabolic activity within and across the regions of interest, that differentiate HIV+ children from HIV- children at 11 years.

5.1 Differences in metabolite concentrations between HIV+ and HU children

5.1.1 Elevated tCho in the BG and MFGM of HIV+ children

In GM-containing regions we detected elevated tCho levels in HIV+ children compared to HU children, as has been observed in earlier studies of HIV in adults - usually associated with an inflammatory response, excessive cell proliferation or an influx of cells (Valcour et al., 2012; Harezlak et al., 2011). This finding of elevated tCho was supported by elevations in the tCho/tCr ratio in the BG and MFGM regions. Neuroinflammation, during HIV infection, is driven by the release of host cytokines and chemokines by infected macrophages, and in the presence of viral proteins, such as Tat (Valcour et al., 2012; Toborek et al., 2003; Persidsky et al., 2000). Previous work has, however, also provided conflicting evidence, suggesting that higher viral loads in children are associated with reduced Cho in the BG (Keller et al., 2004).

In a study by Harezlak et al. (2011), several HIV+ adults were diagnosed with ADC, and elevated Cho/Cr was found to be accompanied by an elevation in the Ins/Cr ratio (Harezlak et al., 2011). It is believed that elevated Cho alone is not enough to conclude that inflammation is occurring. During inflammation, an elevation in Ins is expected to occur together with elevated Cho (Harezlak et al., 2011). Interestingly, elevated absolute Ins accompanied elevated tCho in the MFGM of HIV+ children

at 11 years. Thus, we may have more evidence for inflammation in this region following HIV infection. However, an elevation in the Ins/tCr ratio was not seen in the MFGM.

Affirmation of these findings was attained by further comparison between HIV+ children and a group consisting of all HIV- children - combining HEU and HU children into a single uninfected group. In comparison to all HIV- children, higher Cho levels were still observed in the BG and MFGM of HIV+ children. Further, HIV+ children expressed elevated Ins in the MFGM, with a trend for increased Ins in the BG.

Negative associations, although small, observed between CD4 count at scan, and Ins and tCho concentrations in the MFGM, indicate that HIV+ children with lower CD4 counts have higher MFGM Ins and tCho. This provides valuable insight, revealing that, in addition to HIV+ children expressing higher Ins and tCho than HU controls, the levels of these metabolites are greater in HIV+ children who present with worse clinical conditions at 11 years. Similarly, in HIV+ adults, a negative correlation between nadir CD4 count and Cho/Cr was formerly reported (Winston et al., 2012).

A recent study by Snider et al. (2018) presented a mechanism by which macrophages take up Cho and are activated, leading to greater secretion of cytokines when more Cho is present (Snider et al., 2018). Thus, a link between Cho and macrophage-driven inflammation has been established. Additionally, an accumulation of macrophages and microglia in the BG and frontal WM during HIV infection, has been shown, with these immune cells being implicated in inflammation and less severe forms of HIV-associated neurocognitive disorders (Tavazzi et al., 2014).

In previous analysis in a cohort of children from this study, elevated BG Cho levels were detected in HIV+ children at the age of 5 years, who had been placed on treatment within 12 weeks, compared to those who initiated treatment later, as well as to HIV- children (Mbugua et al., 2016). Moreover, specifically in the HIV+ group, the ratio of CD4/CD8 cells at enrollment was associated with Cho levels at 5 years (Mbugua et al., 2016). This complemented a parallel structural MRI study, in which HIV+ children had larger left and right putamina, left globus pallidus and right nucleus accumbens, despite being placed on ART within the first 18 months of life (Randall et al., 2017). These findings collectively suggested to the authors that elevated BG Cho levels in HIV+ children, were associated with greater cell proliferation and influx in this region at 5 years (Randall et al., 2017).

This appears to contradict our finding of a negative association between CD4 count at scan and MFGM Cho concentration at 11 years. The inconsistency in these results may highlight the differing immune responses at enrollment and at 11 years in HIV+ children. While a lower CD4 count at scan is indicative of HIV-driven CD4 T cell destruction, a lower CD4/CD8 ratio at enrollment does not necessarily imply a

reduction in CD4 T cells. A greater CD8 T cell response has been shown to occur during acute infection, in response to replication of the virus and the host-driven inflammatory response (Bastidas et al., 2014; Freel et al., 2012). Thus, a lower CD4/CD8 ratio at enrollment may ultimately be the result of a stronger CD8 response mounted. Children who mounted a strong CD8 response at enrollment, and therefore had a lower CD4/CD8 ratio, would be expected to experience less inflammation and consequently have lower BG Cho at 5 years.

At 7 years, by contrast, the right putamen volume was found to be lower in HIV+ children compared to HIV- children (Nwosu et al., 2018). The corresponding metabolic analysis at 7 years found no significant differences in BG metabolite concentrations between HIV+ and HU children (Robertson et al., 2018). Thus, the absence of elevated Cho in the BG of HIV+ children, together with decreased right putamen volume might indicate that the initial inflammatory response in the BG had subsided by 7 years.

Elevated Cho in the MFGM has been observed in previous MRS studies of HIV+ adults and children alike (Harezlak et al., 2011; Prado et al., 2011). At 7 years, a subset of these HIV+ children from the CHER trial had elevated Cho levels in the MFGM, but by the age of 9 years there was no difference in Cho concentration in HIV+ and HEU children, relative to the controls (unpublished data).

As such, we had hypothesized that by 11 years the levels of Cho in the MFGM would be no different between HIV+ children and controls. The unexpected observation of elevated Cho might indicate that at 11 years HIV+ children return to a similar inflammatory state to that observed at 7 years. However, it should be noted that at 7 years there was a larger number of HU controls (n=27) than at 9 years (n=18, unpublished data). Thus, although more HIV+ children contributed data at 9 years, there may have been less statistical power to detect differences in Cho levels because of the smaller HU control group and, therefore, elevated Cho in the MFGM of HIV+ children may in fact have been consistent across the ages at which the children were scanned - despite the appearance of temporarily restored Cho levels at 9 years.

5.1.2 Restored NAA and Glutamate levels in the BG and MFGM of HIV+ children

At the age of 11 years, altered levels of Glu and NAA were no longer observed in the BG and MFGM of HIV+ children, contrary to our original hypothesis. At 9 years, a subset of HIV+ children from this study were found to have lower NAA and Glu in the BG, which was suggested to be an indication of neuronal damage and the resulting reductions in neuronal signaling (Robertson et al., 2018). In the MFGM Glu levels alone were found to be reduced at 9 years (unpublished data). This finding in the MFGM could

not result from neuronal damage, as reductions in NAA would otherwise accompany the drop in Glu. It was postulated, rather, that a reduction in Glu alone indicates that glutamate/glutamine cyclical activity is being perturbed in some way, in the MFGM of children with HIV (unpublished). The absence of differences in NAA and Glu at 11 years may imply that neuronal damage and disruptions in glutamatergic neuron signaling are remedied by the age of 11 years in perinatally infected children placed on early ART.

Glu in the BG was positively associated with CD4/CD8 at enrollment in children with HIV, indicating that HIV+ children who potentially mounted a greater CD8 response at enrollment, as seen by a lower CD4/CD8 ratio, had lower Glu levels at 11 years. Lower Ins, conversely, was associated with greater CD4 counts at scan. Thus, HIV+ children with lower CD4 counts at scan, and therefore poorer immunological health, had higher BG Ins at 11 years. Previous research in subsets of children from this study, found that a reduction in the CD4/CD8 ratio at enrollment was associated with reduced Cho and NAA in the BG at 5 years, reduced Cho at 7 years (Robertson et al., 2018; Mbugua et al., 2016), or reduced Ins at 9 years (unpublished data).

Greater tCr in the MFGM at 11 years was associated with lower CD4 counts at enrollment. This is surprising, as the protective role of Cr in the CNS is well-established (for review see Gualano et al., 2010) and, thus, children with poorer clinical measures at enrollment would be expected to have lower Cr levels. Interestingly, a negative correlation between CD4% at enrollment and MFGM Cr concentration at 7 years, was previously found in children from the CHER trial (unpublished data). It should be noted that all the B values calculated for the associations of clinical measures with MFGM metabolites at 11 years, were very small and, consequently, these associations may not have a clinical implication.

5.1.3 Reduction in NAA and Cr in the PWM of HIV+ children

In the WM region selected for this study, reductions in the concentrations of tNAA and tCr were detected in HIV+ children. A reduction in NAA within WM has been observed in other diseases such as multiple sclerosis and schizophrenia, with axonal damage or reduced axonal connectivity as the probable underlying cause (Bjartmar et al., 2001; Lim et al., 1998). In a previous study of HIV+ children and adolescents, the ratio of NAA/Cr in parietal WM was found to be lower in HIV+ individuals with encephalopathy compared to those without encephalopathy, but not compared to controls (Pavlakis et al., 1995). The authors suggested that reduced WM NAA is an indication of axonal injury (Pavlakis et al., 1995). Similarly, our results likely indicate that axonal damage in the PWM is occurring in both HIV+ and HEU children at 11 years. Viral proteins and antiretrovirals, such as gp120 and nucleoside

reverse transcriptase inhibitors, have been shown to cause the shortening of axons (Robinson, Li & Nath, 2007). This is believed to be due mitochondrial injury, resulting from toxicity, or due to the build-up of vesicles and proteins as a result of interrupted transport along microtubules (Robinson, Li & Nath, 2007). Therefore, this may explain how axonal damage may occur in children undergoing treatment for HIV infection, or in children perinatally exposed to HIV and ART.

Further, children who initiated ART at a later age had lower NAA in the PWM at 11 years. This result indicates that earlier treatment may minimize long-term damage in HIV+ children, however, the relevance of this result is questionable as it is so small.

White matter damage has previously been reported in this cohort: diffusion tensor imaging studies at the ages of 5 and 7 years found greater radial diffusivity and lower fractional anisotropy (FA) in specific brain regions of HIV+ children, which was proposed to indicate depletion of, or damage to, the myelin sheath (Jankiewicz et al., 2017; Ackermann et al., 2016). At 5 years, the right corticospinal tract contained the areas where this was most noticeable (Ackermann et al., 2016), while at 7 years lower FA was reported in the left inferior fronto-occipital and left inferior longitudinal fasciculi (Jankiewicz et al., 2017). These papers support our observations of anomalies in WM due to HIV infection.

As Cr provides an internal reference in many MRS studies, reductions in Cr levels are rarely reported and, thus, few papers discuss what a result like this may imply. PCr provides an important energy source and is catabolized to release Cr and phosphate for ATP generation (Gualano et al., 2010). Cyclical activity, enabling continued ATP generation, occurs with the addition of phosphate to Cr in the mitochondria (Steen et al., 2010). Therefore, reductions in tCr in the WM in HIV infection, might indicate a reduction in energy production or reduced energy requirements.

A protective role for Cr in the brain has been suggested, through studies investigating the efficacy of providing Cr supplements to patients with Huntington's disease and Parkinson's disease (Gualano et al., 2010; Bender et al., 2006; Hersch et al., 2006; Dedeoglu et al., 2003). Although impaired mitochondrial activity and energy deficits are implicated in these conditions, the exact method by which Cr supplements provide protection, is unclear. There are suggestions that Cr supplementation may minimise oxidative stress, reduce neuronal damage and improve cognitive outcomes (for review see Gualano et al., 2010), alluding to the possibility that there may be functional and cognitive implications resulting from reduced Cr in children with HIV.

It is possible that the observation of lowered tCr in HIV+ children is related to treatment, rather than HIV-infection. ART drugs, particularly nucleoside reverse transcriptase inhibitors, have been shown to have toxic effects in mitochondria, causing increasing mutations in mitochondrial DNA and, therefore

impacting respiratory chain activity (Ouyang et al., 2018; Payne et al., 2011; Gingelmaier et al., 2009). It has been suggested that ultimately this occurs due to more rapid DNA replication, leading to more errors in the DNA code (Payne et al., 2011). Due to the role of PCr and Cr in ATP generation within the mitochondria (Gualano et al., 2010) a disruption of respiratory chain activity, as a result of mitochondrial damage, would lead to a reduction in tCr.

The ratios of metabolites to Cr, in the PWM, were similar in HIV+ children and HU controls at 11 years. While a trend for elevated Cho/Cr was observed in HIV+ compared to HU children, no such trend in the absolute Cho concentration was detected. In WM brain regions, various studies previously found elevated Cho/Cr in HIV+ children (Van Dalen et al., 2016; Prado et al., 2011). By contrast, in a study by Keller et al. (2004), levels of Cho in the frontal WM were found to be lower for HIV+ children.

The absence of differences in ratios to tCr in this region, given that absolute tNAA and Ins concentrations differed between HIV+ and HU children, are most likely a consequence of the differing levels of Cr between these groups. Thus, if absolute concentrations of metabolites were not considered, important findings could be overlooked. Consequently, although Cr is commonly used as a reference metabolite in 1H-MRS studies, these findings indicate that it may not be a valid reference for studying the effects of HIV within this region of the brain.

It should be noted that a trend for lower Ins was also observed in the PWM of HIV+ children. In terms of determining the functional significance of lower WM Ins, reduced Ins/tCr was previously detected in the frontal lobe of individuals diagnosed with depression (Coupland et al., 2005; Frey et al., 1998) and could have resulted from reductions in glial cells, either due to lower proliferation rates or shorter cell lifespans (Coupland et al., 2005). However, no conclusive information can be drawn from this result at 11 years.

5.2 Inter- and intra-regional metabolite couplings

As previously discussed, there is a cyclical pattern to the expression of several of the metabolites of interest (Steen et al., 2010; Bak, Schousboe & Waagepetersen, 2006; Daikhin & Yudkoff, 2000) and they do not exist independently of each other. L-glutamate, for example, is added to NAA to produce NAAG (for review see Baslow, 2000).

Pearson's correlation matrices provide a means of determining the raw correlations between all the metabolite concentrations within the regions of interest, and across these regions. By comparing correlations generated separately for HIV+ and HIV- children, altered patterns of metabolic activity resulting from HIV infection can be identified. It should be noted in advance that correlations do not

imply that there is a causal relationship between metabolites, but these correlations do provide interesting insight into the relationships between pairs of metabolites.

Spearman's rank correlation analysis was previously used to look at the relationship between different metabolites within the BG of HIV+ adults (Yiannoutsos et al., 2004). A positive correlation was observed between the ratios of Cho/Cr and NAA/Cr, as well as between absolute measures of Cho and NAA in the BG (Yiannoutsos et al., 2004).

We identified many strong within-region metabolite correlations for both HIV+ and HIV- children at 11 years. There were fewer strong correlations between metabolites in different regions of interest and, when comparing the analysis between the two HIV status groups, only a few of these correlations were maintained in both groups with more moderate strength. This supports the findings of an adult MRS study in which strong within-region correlations between metabolites were found, while the across-region correlations between metabolites were negative (Grachev & Apkarian, 2002). This led Grachev et al. (2002) to conclude that the best means of grouping metabolites is according to their voxels. The interesting and more informative correlations, however, are the inter-regional correlations that were strong in HIV+ children yet weak in HIV- children, or vice-versa. Therefore, these correlations were investigated further.

Firstly, the correlation between BG Cho and MFGM Cho was strongly positive in HIV+ children, yet weak in HIV- children. This compliments our linear regression analysis findings, where Cho levels were found to be significantly higher in the BG and MFGM regions of HIV+ children, than in the HU controls. Additionally, this may indicate that there is a common process which occurs in HIV+ children, that is driving the simultaneous rise in Cho in both regions. This finding may suggest that HIV and/or ART are driving inflammation in the BG and MFGM regions of HIV+ children concurrently. In HIV- children however, there is no indication of inflammation.

By contrast, the correlation between BG Cr and PWM Cr was strong and positive in HIV- children, yet weak in HIV+ children. A strong correlation between Cr in these regions, might imply that energy production is strongly coupled in these regions in HIV- children, while there is no clear relationship between energy production in these two regions in children infected with the virus. This result, too, supports our linear regression analysis in which HIV+ children were found to have a reduction in tCr in the PWM region, while no reduction in tCr in the BG region was observed between HIV+ and HU children. Thus, the reduction in tCr in the PWM in HIV+ children is not accompanied by or coupled to a reduction in tCr in the BG.

Strong, positive pairwise correlations between Ins levels in the 3 different voxels, were observed in HIV- children. These correlations, though not as strong, were still of intermediate strength in HIV+ children. This finding is interesting, as HIV+ children had elevated absolute Ins in the MFGM, a trend for elevated Ins in the BG, but contrastingly they had a trend for reduced Ins in the PWM compared to HU children. Thus, one might expect a negative correlation between Ins in the PWM and Ins in the other regions. The result we obtained, however, indicates that children with higher Ins in one region tend to have higher Ins in the other regions, and this is true of both HIV+ and HIV- children - although the pattern is more evident in HIV- children.

Although this approach of using correlation matrices for studying metabolic differences between HIV+ and HIV- children is informative, it requires a manual approach for comparing the groups. Moreover, it does not take into account other contributing factors or confounding variables and it does not identify neurometabolic patterns between more than two metabolites, which can distinguish between the two HIV status groups. Thus, it is more likely that important changes in metabolic activity, driven by HIV infection, may be masked by other factors and, therefore, go unnoticed when using this approach.

A previous study by Van Dalen et al. (2016) in children, showed strong associations between the levels of several individual metabolites across GM and WM, as well as relationships between metabolites within GM and WM when adjusting for HIV status (Van Dalen et al., 2016). A study in adults using J-difference spectroscopy, additionally found several intra-regional metabolite associations, yet no inter-regional metabolite correlations except for Glu in healthy individuals (Waddell et al., 2011). Thus, it is interesting that we noted numerous cross-regional metabolite correlations in uninfected children at 11 years. We cannot conclude from these findings whether there is a direct or indirect relationship between metabolite concentrations in the different regions of interest. However, it is clear that metabolic activity in the different regions of the brain are not independent.

5.3 Metabolic factors that distinguish HIV+ from HIV- children

The linear regression approach presented earlier, does not consider the inter-regional covariations of a particular metabolite or the relationships between different metabolites within a region. However, there is reason to expect that response patterns of multiple metabolites, measured across different regions of the brain, would be related due to complex underlying networks of metabolic pathways (for review see Rae, 2014) - which may not be directly observed. Although Pearson's correlation analysis is somewhat able to explore these patterns in terms of pairwise comparisons, there is more scope for collectively investigating the relationships between numerous metabolites.

Factor analysis reduces a large number of correlated variables into a smaller number of latent variables or factors, enabling interesting patterns of metabolic activity to be identified. Factor scores can be incorporated into logistic regression models to determine if particular metabolite combinations can be used to distinguish one group of individuals from another.

A study by Yiannoutsos et al. (2004) utilized factor analysis, followed by logistic regression, as a mechanism to identify patterns of metabolite expression that could be used to differentiate between HIV+ adults with ADC and those without symptoms of dementia. This approach is advantageous because confounding variables can be adjusted for, and ROC curves can then provide measures of the predictive capabilities of the models generated.

Adapting the approach of Yiannoutsos et al. (2004), we assessed differences in metabolic patterns in HIV+ vs HIV- children. Initial investigative factor analysis with 5 metabolites across the 3 regions of interest, led to factors with very region-specific combinations of metabolites. This would be expected due to the strong intra-regional correlations between metabolites, as shown in the Pearson's correlation analysis. However, patterns of inter-regional correlations were not seen as a result.

Narrowing down the analysis to 3 metabolites across the 3 regions, we identified 5 factors representing patterns of metabolic activity. Surprisingly, only the inflammatory factor grouped a single metabolite, tCho, across all regions. However, NAA and Cr were grouped within the MFGM, as well as within the PWM, reflecting an association between these metabolites within both of these regions. Although the BG neuronal factor was driven largely by BG NAA, substantial loadings were also contributed by NAA in the MFGM and PWM, as well as BG tCho and tCr.

Due to the exploratory nature of this work, both *varimax* (orthogonal) and *promax* (oblique) rotations were used (Finch, 2006; Kaiser, 1958). Both approaches yielded similar weighted factor scores, while simple scores were identical for these approaches.

5.3.1 The multiregional inflammatory factor

The introduction of weighted factor scores, based on *varimax* rotation, into a logistic regression model, showed that the inflammatory factor - driven by tCho in the 3 regions of interest - and the PWM axonal factor - driven by PWM tCr and tNAA - distinguished between HIV+ and HIV- children at 11 years. A similar conclusion was obtained when introducing the weighted scores based on an oblique *promax* rotation (Finch, 2006; Abdi, 2003) and when using a simple scoring approach (as in Yiannoutsos et al., 2004).

HIV+ children had higher inflammatory factor scores, than HIV- children, because of higher combined concentrations of tCho in the 3 regions of interest. This result reinforces our linear regression analysis findings and confirms that the correlation between Cho in the BG and MFGM is strong in HIV+ children and weak in HIV- children, as shown by Pearson's correlation analysis. Yet, it also implicates elevated WM Cho in HIV infection and adds another dimension to our previous analysis. The nature of this pattern, in terms of whether there is a mutual underlying metabolic pathway, or a causal relationship responsible for the simultaneous elevation in Cho, cannot be determined from this analysis alone.

Cross-regional Cho elevation may indicate a simultaneous inflammatory response across the brain, in both gray and white matter. Several paediatric studies have shown that children and adolescents suffering from depression, express higher levels of Cho in the frontal lobe (MacMaster & Kusumakar, 2006; Farchione, Moore & Rosenberg, 2002; Steingard et al., 2000) and the left caudate (Gabbay et al., 2007). This raises the possibility that elevated Cho in HIV+ children may be the manifestation of - or potentially play a mechanistic role in - mental health disorders associated with HIV infection. However, further investigation would be required to determine if the HIV+ children in this study are experiencing depression or other mental health issues.

Related work using spectroscopic imaging similarly found a Cho factor, reflecting elevated Cho across deep gray and WM regions, that distinguished between HIV+ patients and controls (Mohamed, Lentz et al., 2010). However, this factor did not differentiate between HIV+ adults with and without dementia (Mohamed, Lentz et al., 2010). The work by Yiannoutsos et al. (2004) in adults, interestingly, also found that an inflammatory factor, driven by Cho/Cr and Ins/Cr ratios in parietal cortex and centrum semiovale voxels, did not distinguish between individuals with and without dementia. However, a factor representing elevated NAA/Cr and Cho/Cr in the BG, as well as a factor dependent on NAA/Cr reductions in centrum semiovale and parietal cortex, were strongly associated with dementia (Yiannoutsos et al., 2004). Mohamed et al (2010) similarly found that a factor dependent on NAA in the frontal WM, parietal WM, centrum semiovale and parietal GM, distinguished between HIV+ subjects with and without dementia (Mohamed, Lentz et al., 2010). Thus, it may be interesting to look at whether the factors generated in this study at 11 years could be useful in investigating HIV-associated neurocognitive disorders or delayed neurodevelopment in these children.

As pairwise correlations between Ins in the different voxels were observed through Pearson's correlation analysis, it would be interesting to investigate whether a cross-regional Ins factor, or factors coupling Ins with Cho, could distinguish between HIV+ and HIV- children.

5.3.2 The white matter axonal factor

In our study of 11-year-old children, lower PWM axonal factor scores, resulting from lower coupled PWM tNAA and tCr concentrations, were found in HIV+ children compared to HIV- children. Again, this reinforces our findings from linear regression analysis, while providing evidence of an association between NAA and Cr in WM.

In a similar way, an early study by Salvan et al. (1998) utilised principal component analysis to assess metabolic patterns in the centrum semiovale of HIV+ children between 2-10 years. Reduced NAA, elevated Ins and elevated lipid, responsible for the first axis, separated HIV+ children with encephalopathy from control children and from HIV+ children without encephalopathy, indicating that WM damage had occurred (Salvan et al., 1998).

Since many studies considered ratios to Cr, including the factor analysis study carried out by Yiannoutsos et al. (2004) and the study by Van Dalen et al. (2016), a within-region association of NAA and Cr has not to our knowledge been identified before now.

5.3.3 Predictive ability of metabolic factors

The predictive ability of our logistic regression models, incorporating factor scores, appear to be relatively strong given AUC percentages above 75%. The logistic regression model with the greatest predictive ability contained the weighted scores from factor analysis with *varimax* rotation. Whilst an MRS approach would not necessarily be a good stand-alone diagnostic tool for HIV infection, as metabolic changes can be the outcome of numerous different pathologies, utilising factor analysis and logistic regression can enable the identification of biomarkers of HIV infection and highlight metabolite combinations that are altered with infection. Nevertheless, it would be interesting to investigate the potential in this approach for tracking the prognosis of HIV+ individuals, in particular the neurometabolic impacts with increasing disease severity.

5.4 Differences in metabolic activity between HEU and HU children

5.4.1 Restored metabolite levels in the BG of HEU children

Based on the MRS study of the BG region at 9 years, carried out by Robertson et al. (2018), it was hypothesised that reductions in Cho and Cr might occur at 11 years in the BG region of HEU children. However, no meaningful differences in BG metabolite levels were detected in comparison to the control children. This indicates that by 11 years, the harmful effects of perinatal HIV exposure are no longer evident in the BG of HEU children.

5.4.2 Lower Glx and Cr in the MFGM in HEU children

Although it was hypothesized, based on the findings at 9 years (unpublished data), that HEU children would have lower MFGM Cho levels at 11 years, this was not observed. However, differences in other metabolite levels were observed for 11-year-old HEU children in comparison to HU children. HEU children were found to have lower absolute Glx concentrations and a reduction in absolute Cr in the MFGM. However, no reduction in Glx/tCr was observed, likely due to simultaneous reductions in Glx and Cr in these children.

Lower MFGM Glu was previously reported in a subset of HIV+ children from this study when they were 7 and 9 years old (unpublished data). HEU children were also found to have lower Glu at 9 years, but this was specific to the BG region (Robertson et al., 2018). A reduction in the ratio of Glu to tCr within GM has previously been reported to associate with lower nadir CD4 counts in HIV+ children (Van Dalen et al., 2016).

Reduced Glu has similarly been reported in frontal GM of HIV+ adults with HIV-associated neurocognitive disorders (Ernst et al., 2010). Several possible reasons for lower Glu in the frontal GM have been suggested, including a disruption in the uptake of Glu by astrocytes, leading to a build-up of Glu in extracellular space and, ultimately, neurotoxicity (Ernst et al., 2010). Reduced NAA was additionally observed by Ernst et al. (2010) in the HIV+ adults, which indicated putative neuronal damage. In the HEU children in this study at 11 years, however, no reductions in NAA in the MFGM were seen, indicating that this is an unlikely scenario in these children.

A second possibility presented by Ernst et al. (2010), was that Glu may provide the amino acid components required in the repair of damaged cell membranes (Ernst et al., 2010). If this were the case at 11 years, however, one would expect to observe a corresponding elevation in Cho, which is not seen in HEU children.

Finally, it was also proposed that hampered mitochondrial activity could lead to a reduction in Glu (Ernst et al., 2010). The anapleurotic reactions involved in Glu synthesis via the tricarboxylic acid cycle, occur within the mitochondria (for review see Sonnewald, 2014). Thus, mitochondrial damage could lead to the disruption of Glu synthesis. However, in addition to observing lower Glu, damage to mitochondria would be expected to result in a reduction in Cr levels and, ultimately, lower energy production (for review see Gualano et al., 2010). Interestingly, a reduction in the concentration of Cr was also observed in HEU children, suggesting that this is the most probable scenario occurring in HEU children at 11 years.

The reduction we observed in the MFGM was specifically in Glx levels, not Glu alone, suggesting that whatever process is being affected in HEU children, it is impacting combined Glu and Gln levels. Due to the cyclical activity of these metabolites (Bak, Schousboe & Waagepetersen, 2006), it is possible that the synthesis of Glu is being affected at a mitochondrial level, as this would eventually also lead to less Gln being present. However, as Glu is the main contributor to Glx levels (Mohamed, Barker et al., 2010), it is surprising to not see at least a trend in reduced Glu in HEU children -although this may be due to slightly improved accuracy when measuring Glx levels, compared to Glu. So, differences in Glu levels between HIV status groups may be less distinguishable.

5.4.3 Reduced tNAA in the PWM of HEU children

Lower concentrations of tNAA occurred in the PWM of HEU children in comparison to HU children. This reveals that the effects of perinatal exposure to HIV, persist at 11 years in this region. As the PWM is a region of late myelination (Parazzini et al., 2002), the effects of HIV infection and exposure may only occur at a later stage, which could explain why the impact of exposure is observed in this region at 11 years for HEU children, while metabolic activity in the BG is restored.

HEU children were found to have no clear differences in the ratios of the various metabolites to Cr in the PWM at 11 years, compared to HU controls.

5.4.4 Investigating metabolic patterns in HEU children

In addition to comparing factor scores between HIV+ and HIV- children, changes in metabolic patterns in response to HIV exposure, may be of interest. There was not enough power to investigate whether factor scores could distinguish between HEU and HU children when using logistic regression. Linear regression analysis could, however, be utilized to compare factor scores between our initial 3 groups of interest. Unfortunately, this meant that the predictive strength of our models could not be evaluated.

This analysis provided additional insight, showing that levels of tCr and tNAA in the PWM were also lower in HEU children, in addition to HIV+ children having reductions in these metabolite measures. It is interesting that combining HEU and HU children into a single uninfected group, did not prevent the reduction in PWM tCr and tNAA in HIV+ children from being observed. Unlike HIV+ children, HEU children did not have elevated multiregional inflammatory factor scores compared to controls. Therefore, there is no indication of inflammation as a result of perinatal exposure to HIV and ART.

5.5 Limitations

Although this study provides insight into the metabolic changes occurring in perinatally infected children, a limitation is that we are unable to differentiate between the effects of HIV and ART. At this point we have not considered the possibility that other treatment being provided, such as for attention deficit hyperactivity disorder (ADHD) may alter neurometabolic activity in children (Husarova et al., 2014; Wiguna et al., 2012). As previously mentioned, depression alters metabolic activity (Gabbay et al., 2007; Farchione, Moore & Rosenberg, 2002; Steingard et al., 2000) and thus, diagnoses for depression and ADHD, could be considered in future studies particularly as this cohort progress to adolescence.

This study is cross-sectional and, thus, only compares metabolic differences between the three HIV status groups within a single age bracket. Longitudinal analysis of changes in metabolic activity over time, would provide valuable insight into neurological effects with the progression of disease and with continued exposure to ART in HIV+ children undergoing treatment.

In this study we point out the issues of using tCr as a reference metabolite for studying metabolic changes associated with perinatal HIV infection in children. It should also be pointed out that although we presented analysis of 'absolute' metabolite concentrations, these measures also rely on an internal reference – in this case water. This depends on the assumption that we know the concentration of water in WM and GM. Thus, if the water concentration in these tissues were to be impacted by HIV infection, our measures of metabolite concentrations would be incorrectly estimated.

A further limitation is that metabolite levels in the brain may be affected by the circadian rhythm (Arm et al., 2019; Naismith et al., 2014). Thus, the time of day at which scanning is carried out may impact the spectra obtained and, in future, could be considered. However, it is impractical to account for all possible sources of variation between study participants in regression analysis, and thus we seek to develop models which are useful and through which we can attempt to draw meaningful conclusions.

Relationships between viral load and metabolite concentrations have not been examined at younger ages in cohorts from this study, and this may be due to viral load being somewhat uninformative, as there are upper and lower thresholds of detection. Additionally, in this study at 11 years, viral load was categorized in such a way that only 2 children had measurable viral load and so we were unable to investigate the relationship between viral load at scan and metabolite concentrations.

Finally, clinical data collected at 11 years was excluded if it was obtained more than 6 months from the date of scan. This is a limitation as it does not provide an accurate indication of the health of participants at the time of scan. Ideally, clinical data collected with 1 month from the date of scan should be used, however, in doing so, there would have been insufficient data to carry out this analysis.

5.6 Future work

As this work contributes to a larger, longitudinal study, further work could be done to assess how the patterns of metabolic activity identified through factor analysis, change over time in children perinatally infected with or exposed to HIV and ART.

This research could potentially provide a platform for comparing ART regimens in future, to assess which antiretroviral combinations have the least harmful or toxic effects. Further, as a cohort of South African children have been enrolled in this study, this work may also present a set of neurometabolic profile references, specific to our population.

As macrophage-driven cytokine secretion is largely responsible for neuroinflammation during HIV infection (Tavazzi et al., 2014), and we observe signs of inflammation in 11 year-old HIV+ children at a metabolic level, an exploration of the relationship between metabolite and cytokine responses could be very informative. In particular, it would be interesting to investigate whether inflammation is occurring due to the inability to maintain viral suppression. Further, we could consider what proportion of HIV+ children had early treatment or deferred treatment, and whether inflammation could be the result of viral escape.

Finally, various studies in adults have sought to tie together the findings of MRS analysis and other modalities. A study by Ernst et al. (2003), utilising fMRI and MRS modalities to study HIV infection in adults, showed that greater BOLD signaling within the memory network of HIV+ adults, correlated with elevated glial metabolites (Cho, Ins and Cr) during HIV infection prior to dementia. The authors suggested that this finding is an indication of inflammation and glial cell proliferation being largely responsible for the damage seen at this stage of HIV infection (Ernst, Chang & Arnold, 2003).

Paul et al. (2008) investigated how MRS and structural measures linked to cognitive changes occurring in HIV+ adults without dementia. In this study they showed correlations between cognitive performance measures and the size of the caudate structure, as well as Ins/Cr, NAA/Ins and NAA/Cho metabolite levels (Paul et al., 2008).

It would be interesting to investigate similar questions in children at 11 years. Definitive conclusions regarding the effects of HIV during this stage of brain development cannot be drawn based on MRS analysis alone. Therefore, additional study could seek to tie in parallel structural and functional studies at 11 years, with these spectroscopic findings. This inter-disciplinary approach may provide a more holistic picture of the continued effects of perinatal HIV infection and early ART, as children enter into adolescence - providing more insight into what these metabolic changes mean in terms of alterations to signalling pathways, structural changes and cognitive impacts of HIV and early treatment on brain development.

5.7 Conclusion

This study provides a profile for the differences in metabolite concentrations and patterns between HIV+ children on treatment, HEU children and HU controls at 11 years. The majority of MRS studies focusing on HIV in children, have been carried out in economically developed countries (Van Dalen et al., 2016; Nagarajan et al., 2012; Gabis et al., 2006; Keller et al., 2004; Pavlakis et al., 1995). This study, however, provides findings specific to a cohort of South African children, from a resource-limited setting and in a country where the burden of HIV is particularly high.

Based on the findings of this study, the neurometabolic effects of perinatal HIV infection seen at 11 years include elevated tCho in the BG and MFGM regions, and lower tNAA and tCr in the PWM. Differing tCr levels, between HIV+ and HU children, suggest that this may not be a reliable internal reference metabolite.

These results indicate inflammation in GM and axonal damage in WM, despite HIV+ children beginning early treatment. However, whether these results can be extrapolated to gray and white matter across the entire brain, or whether these results are specific solely to the regions we assessed, cannot be determined from this study alone. Finally, the implications of these metabolic alterations at a cognitive level, need to be investigated further.

6. References

Abdi, H. 2003. Factor rotations in factor analyses. *Encyclopedia for Research Methods for the Social Sciences.Sage: Thousand Oaks, CA.* :792-795.

Ackermann, C., Andronikou, S., Laughton, B., Kidd, M., Dobbels, E., Innes, S., van Toorn, R. & Cotton, M. 2014. White matter signal abnormalities in children with suspected HIV-related neurologic disease on early combination antiretroviral therapy. *The Pediatric Infectious Disease Journal.* 33(8):e207.

Ackermann, C., Andronikou, S., Saleh, M.G., Laughton, B., Alhamud, A.A., van der Kouwe, A., Kidd, M., Cotton, M.F. et al. 2016. Early antiretroviral therapy in HIV-infected children is associated with diffuse white matter structural abnormality and corpus callosum sparing. *American Journal of Neuroradiology.* 37(12):2363-2369.

Alexander, G.E., DeLong, M.R. & Strick, P.L. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience.* 9(1):357-381.

Alimonti, J.B., Ball, T.B. & Fowke, K.R. 2003. Mechanisms of CD4 T lymphocyte cell death in human immunodeficiency virus infection and AIDS. *Journal of General Virology.* 84(7):1649-1661.

Allen, T.M., Altfeld, M., Geer, S.C., Kalife, E.T., Moore, C., O'Sullivan, K.M., DeSouza, I., Feeney, M.E. et al. 2005. Selective escape from CD8 T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. *Journal of Virology.* 79(21):13239-13249.

Anderson, E., Zink, W., Xiong, H. & Gendelman, H.E. 2002. HIV-1-associated dementia: a metabolic encephalopathy perpetrated by virus-infected and immune-competent mononuclear phagocytes. *Journal of Acquired Immune Deficiency Syndromes (1999).* 31:43.

Anthony, I.C. & Bell, P.J. 2008. The neuropathology of HIV/AIDS. *International Review of Psychiatry.* 20(1):15-24.

Anthony, I.C., Ramage, S.N., Carnie, F.W., Simmonds, P. & Bell, J.E. 2005. Influence of HAART on HIV-related CNS disease and neuroinflammation. *Journal of Neuropathology & Experimental Neurology.* 64(6):529-536.

Arm, J., Al-iedani, O., Lea, R., Lechner-Scott, J. & Ramadan, S. 2019. Diurnal variability of cerebral metabolites in healthy human brain with 2D localized correlation spectroscopy (2D L-COSY). *Journal of Magnetic Resonance Imaging.*

Bak, L.K., Schousboe, A. & Waagepetersen, H.S. 2006. The glutamate/GABA-glutamine cycle: aspects of transport, neurotransmitter homeostasis and ammonia transfer. *Journal of Neurochemistry.* 98(3):641-653.

Bangsberg, D.R., Charlebois, E.D., Grant, R.M., Holodniy, M., Deeks, S.G., Perry, S., Conroy, K.N., Clark, R. et al. 2003. High levels of adherence do not prevent accumulation of HIV drug resistance mutations. *Aids.* 17(13):1925-1932.

Barger, S.W., Goodwin, M.E., Porter, M.M. & Beggs, M.L. 2007. Glutamate release from activated microglia requires the oxidative burst and lipid peroxidation. *Journal of Neurochemistry.* 101(5):1205-1213.

Baslow, M.H. 2000. Functions of N-acetyl-l-aspartate and N-acetyl-l-aspartylglutamate in the vertebrate brain: role in glial cell-specific signaling. *Journal of Neurochemistry.* 75(2):453-459.

Baslow, M.H. 2010. Evidence that the tri-cellular metabolism of N-acetylaspartate functions as the brain's "operating system": how NAA metabolism supports meaningful intercellular frequency-encoded communications. *Amino Acids.* 39(5):1139-1145.

Bastidas, S., Graw, F., Smith, M.Z., Kuster, H., Günthard, H.F. & Oxenius, A. 2014. CD8 T cells are activated in an antigen-independent manner in HIV-infected individuals. *The Journal of Immunology.* 192(4):1732-1744.

Bender, A., Koch, W., Elstner, M., Schombacher, Y., Bender, J., Moeschl, M., Gekeler, F., Müller-Myhsok, B. et al. 2006. Creatine supplementation in Parkinson disease: a placebo-controlled randomized pilot trial. *Neurology.* 67(7):1262-1264.

Berger, E.A., Murphy, P.M. & Farber, J.M. 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annual Review of Immunology.* 17(1):657-700.

Berger, J.R. & Arendt, G. 2000. HIV dementia: the role of the basal ganglia and dopaminergic systems. *Journal of Psychopharmacology.* 14(3):214-221.

Bjartmar, C., Kinkel, R.P., Kidd, G., Rudick, R.A. & Trapp, B.D. 2001. Axonal loss in normal-appearing white matter in a patient with acute MS. *Neurology.* 57(7):1248-1252.

Bonelli, R.M. & Cummings, J.L. 2007. Frontal-subcortical circuitry and behavior. *Dialogues in Clinical Neuroscience.* 9(2):141.

Booth, J.R., Wood, L., Lu, D., Houk, J.C. & Bitan, T. 2007. The role of the basal ganglia and cerebellum in language processing. *Brain Research.* 1133:136-144.

Braissant, O., Béard, E., Torrent, C. & Henry, H. 2010. Dissociation of AGAT, GAMT and SLC6A8 in CNS: relevance to creatine deficiency syndromes. *Neurobiology of Disease.* 37(2):423-433.

Braissant, O., Henry, H., Béard, E. & Uldry, J. 2011. Creatine deficiency syndromes and the importance of creatine synthesis in the brain. *Amino Acids.* 40(5):1315-1324.

Brodt, H.R., Kamps, B.S., Gute, P., Knupp, B., Staszewski, S. & Helm, E.B. 1997. Changing incidence of AIDS-defining illnesses in the era of antiretroviral combination therapy. *Aids.* 11(14):1731-1738.

Casey, B.J., Jones, R.M. & Hare, T.A. 2008. The adolescent brain. *Annals of the New York Academy of Sciences.* 1124(1):111-126.

Chang, L., Speck, O., Miller, E.N., Braun, J., Jovicich, J., Koch, C., Itti, L. & Ernst, T. 2001. Neural correlates of attention and working memory deficits in HIV patients. *Neurology.* 57(6):1001-1007.

Chen, Y., An, H., Zhu, H., Stone, T., Smith, J.K., Hall, C., Bullitt, E., Shen, D. et al. 2009. White matter abnormalities revealed by diffusion tensor imaging in non-demented and demented HIV patients. *NeuroImage.* 47(4):1154-1162.

Chen, Z. & Trapp, B.D. 2016. Microglia and neuroprotection. *Journal of Neurochemistry.* 136(S1):10-17.

Churchill, M. & Nath, A. 2013. Where does HIV hide? A focus on the central nervous system. *Current Opinion in HIV and AIDS.* 8(3):165.

Clavel, F. & Hance, A.J. 2004. HIV drug resistance. *New England Journal of Medicine.* 350(10):1023-1035.

Cohen, S., Caan, M.W., Mutsaerts, H., Scherpbier, H.J., Kuijpers, T.W., Reiss, P., Majoie, C.B. & Pajkrt, D. 2016. Cerebral injury in perinatally HIV-infected children compared to matched healthy controls. *Neurology.* 86(1):19-27.

Correa, R. & Munoz-Fernandez, M.A. 2001. Viral phenotype affects the thymic production of new T cells in HIV-1-infected children. *AIDS (London, England).* 15(15):1959-1963.

Cotton, M.F., Violari, A., Otwombe, K., Panchia, R., Dobbels, E., Rabie, H., Josipovic, D., Liberty, A. et al. 2013. Early time-limited antiretroviral therapy versus deferred therapy in South African infants infected with HIV: results from the children with HIV early antiretroviral (CHER) randomised trial. *Lancet.* 382(9904):1555-1563. DOI:10.1016/S0140-6736(13)61409-9 Available: http://www.ncbi.nlm.nih.gov/pubmed/24209829.

Coupland, N.J., Ogilvie, C.J., Hegadoren, K.M., Seres, P., Hanstock, C.C. & Allen, P.S. 2005. Decreased prefrontal Myo-inositol in major depressive disorder. *Biological Psychiatry.* 57(12):1526-1534.

Cysique, L.A. & Brew, B.J. 2009. Neuropsychological functioning and antiretroviral treatment in HIV/AIDS: a review. *Neuropsychology Review.* 19(2):169-185.

Dahiru, T. 2008. P-value, a true test of statistical significance? A cautionary note. *Annals of Ibadan Postgraduate Medicine.* 6(1):21-26.

Dahl, V., Josefsson, L. & Palmer, S. 2010. HIV reservoirs, latency, and reactivation: prospects for eradication. *Antiviral Research.* 85(1):286-294.

Dai, G., Yu, H., Kruse, M., Traynor-Kaplan, A. & Hille, B. 2016. Osmoregulatory inositol transporter SMIT1 modulates electrical activity by adjusting PI (4, 5) P2 levels. *Proceedings of the National Academy of Sciences.* :201606348.

Daikhin, Y. & Yudkoff, M. 2000. Compartmentation of brain glutamate metabolism in neurons and glia. *The Journal of Nutrition.* 130(4):1026S-1031S.

de Almeida, S.M., Rotta, I., Ribeiro, C.E., Smith, D., Wang, R., Judicello, J., Potter, M., Vaida, F. et al. 2016. Blood-CSF barrier and compartmentalization of CNS cellular immune response in HIV infection. *Journal of Neuroimmunology.* 301:41-48.

De Cock, K.M., Fowler, M.G., Mercier, E., De Vincenzi, I., Saba, J., Hoff, E., Alnwick, D.J., Rogers, M. et al. 2000. Prevention of mother-to-child HIV transmission in resource-poor countries: translating research into policy and practice. *Jama.* 283(9):1175-1182.

De Graaf, R.A. 2013. *In vivo NMR spectroscopy: principles and techniques.* John Wiley & Sons. Available: .

Debeaudrap, P., Bodeau-Livinec, F., Pasquier, E., Germanaud, D., Ndiang, S.T., Nlend, A.N., Ndongo, F.A., Guemkam, G. et al. 2018. Neurodevelopmental outcomes in HIV-infected and uninfected African children. *Aids.* 32(18):2749-2757.

Dedeoglu, A., Kubilus, J.K., Yang, L., Ferrante, K.L., Hersch, S.M., Beal, M.F. & Ferrante, A.R.J. 2003. Creatine therapy provides neuroprotection after onset of clinical symptoms in Huntington's disease transgenic mice. *Journal of Neurochemistry.* 85(6):1359-1367.

Derdeyn, C.A. & Silvestri, G. 2005. Viral and host factors in the pathogenesis of HIV infection. *Current Opinion in Immunology.* 17(4):366-373.

Diamond, A. 2002. Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. *Principles of Frontal Lobe Function.* :466-503.

Ernst, T., Chang, L. & Arnold, S. 2003. Increased glial metabolites predict increased working memory network activation in HIV brain injury. *NeuroImage.* 19(4):1686-1693.

Ernst, T., Kreis, R. & Ross, B.D. 1993. Absolute quantitation of water and metabolites in the human brain. I. Compartments and water. *Journal of Magnetic Resonance, Series B.* 102(1):1-8.

Ernst, T., Jiang, C.S., Nakama, H., Buchthal, S. & Chang, L. 2010. Lower brain glutamate is associated with cognitive deficits in HIV patients: A new mechanism for HIV-associated neurocognitive disorder. *Journal of Magnetic Resonance Imaging.* 32(5):1045-1053.

Farchione, T.R., Moore, G.J. & Rosenberg, D.R. 2002. Proton magnetic resonance spectroscopic imaging in pediatric major depression. *Biological Psychiatry.* 52(2):86-92.

Finch, H. 2006. Comparison of the performance of varimax and promax rotations: Factor structure recovery for dichotomous items. *Journal of Educational Measurement.* 43(1):39-52.

Freel, S.A., Picking, R.A., Ferrari, G., Ding, H., Ochsenbauer, C., Kappes, J.C., Kirchherr, J.L., Soderberg, K.A. et al. 2012. Initial HIV-1 antigen-specific CD8 T cells in acute HIV-1 infection inhibit transmitted/founder virus replication. *Journal of Virology.* 86(12):6835-6846.

Frey, R., Metzler, D., Fischer, P., Heiden, A., Scharfetter, J., Moser, E. & Kasper, S. 1998. Myo-inositol in depressive and healthy subjects determined by frontal 1H-magnetic resonance spectroscopy at 1.5 tesla. *Journal of Psychiatric Research.* 32(6):411-420.

Gabbay, V., Hess, D.A., Liu, S., Babb, J.S., Klein, R.G. & Gonen, O. 2007. Lateralized caudate metabolic abnormalities in adolescent major depressive disorder: a proton MR spectroscopy study. *American Journal of Psychiatry.* 164(12):1881-1889.

Gabis, L., Belman, A., Huang, W., Milazzo, M. & Nachman, S. 2006. Clinical and Imaging Study of Human Immunodeficiency Virus-1—Infected Youth Receiving Highly Active Antiretroviral Therapy: Pilot Study Using Magnetic Resonance Spectroscopy. *Journal of Child Neurology.* 21(6):485-490.

Gardner, E.M., Burman, W.J., Steiner, J.F., Anderson, P.L. & Bangsberg, D.R. 2009. Antiretroviral medication adherence and the development of class-specific antiretroviral resistance. *AIDS (London, England).* 23(9):1035.

Gardner, E.M., Sharma, S., Peng, G., Hullsiek, K.H., Burman, W.J., MacArthur, R.D., Chesney, M., Telzak, E.E. et al. 2008. Differential adherence to combination antiretroviral therapy is associated with virological failure with resistance. *AIDS (London, England).* 22(1):75.

Gasparovic, C., Song, T., Devier, D., Bockholt, H.J., Caprihan, A., Mullins, P.G., Posse, S., Jung, R.E. et al. 2006. Use of tissue water as a concentration reference for proton spectroscopic imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine.* 55(6):1219-1226.

George, R., Andronikou, S., Du Plessis, J., du Plessis, A., Van Toorn, R. & Maydell, A. 2009. Central nervous system manifestations of HIV infection in children. *Pediatric Radiology.* 39(6):575-585.

Giedd, J.N. 2004. Structural magnetic resonance imaging of the adolescent brain. *Annals of the New York Academy of Sciences.* 1021(1):77-85.

Gingelmaier, A., Grubert, T.A., Kost, B.P., Setzer, B., Lebrecht, D., Mylonas, I., Mueller-Hoecker, J., Jeschke, U. et al. 2009. Mitochondrial toxicity in HIV type-1-exposed pregnancies in the era of highly active antiretroviral therapy. *Antivir Ther.* 14(3):331-338.

Govindaraju, V., Young, K. & Maudsley, A.A. 2000. Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance in Vivo.* 13(3):129-153.

Grachev, I.D. & Apkarian, A.V. 2002. Multi-chemical networking profile of the living human brain: potential relevance to molecular studies of cognition and behavior in normal and diseased brain. *Journal of Neural Transmission.* 109(1):15-33.

Gualano, B., Artioli, G.G., Poortmans, J.R. & Junior, A.H.L. 2010. Exploring the therapeutic role of creatine supplementation. *Amino Acids.* 38(1):31-44.

Haase, A., Frahm, J., Hanicke, W. & Matthaei, D. 1985. 1H NMR chemical shift selective (CHESS) imaging. *Physics in Medicine & Biology.* 30(4):341.

Hamers, R.L., Sigaloff, K.C., Wensing, A.M., Wallis, C.L., Kityo, C., Siwale, M., Mandaliya, K., Ive, P. et al. 2012. Patterns of HIV-1 drug resistance after first-line antiretroviral therapy (ART) failure in 6 sub-Saharan African countries: implications for second-line ART strategies. *Clinical Infectious Diseases.* 54(11):1660-1669.

Harezlak, J., Buchthal, S., Taylor, M., Schifitto, G., Zhong, J., Daar, E.S., Alger, J., Singer, E. et al. 2011. Persistence of hiv− associated cognitive impairment, inflammation and neuronal injury in era of highly active antiretroviral treatment. *AIDS (London, England).* 25(5):625.

Helms, G. 2008. The principles of quantification applied to in vivo proton MR spectroscopy. *European Journal of Radiology.* 67(2):218-229.

Hersch, S.M., Gevorkian, S., Marder, K., Moskowitz, C., Feigin, A., Cox, M., Como, P., Zimmerman, C. et al. 2006. Creatine in Huntington disease is safe, tolerable, bioavailable in brain and reduces serum 8OH2' dG. *Neurology.* 66(2):250-252.

Herting, M.M., Uban, K.A., Williams, P.L., Gautam, P., Huo, Y., Malee, K., Yogev, R., Csernansky, J. et al. 2015. Default mode connectivity in youth with perinatally acquired HIV. *Medicine.* 94(37).

Holmes, M.J., Robertson, F.C., Little, F., Randall, S.R., Cotton, M.F., van der Kouwe, Andre JW, Laughton, B. & Meintjes, E.M. 2017. Longitudinal increases of brain metabolite levels in 5-10 year old children. *PloS One.* 12(7):e0180973.

Horn, P.S., Feng, L., Li, Y. & Pesce, A.J. 2001. Effect of outliers and nonhealthy individuals on reference interval estimation. *Clinical Chemistry.* 47(12):2137-2145.

Howe, F.A. & Opstad, K.S. 2003. 1H MR spectroscopy of brain tumours and masses. *NMR in Biomedicine.* 16(3):123-131.

Husarova, V., Bittsansky, M., Ondrejka, I. & Dobrota, D. 2014. Prefrontal grey and white matter neurometabolite changes after atomoxetine and methylphenidate in children with attention deficit/hyperactivity disorder: A 1H magnetic resonance spectroscopy study. *Psychiatry Research: Neuroimaging.* 222(1-2):75-83.

Jahanshad, N., Couture, M., Prasitsuebsai, W., Nir, T.M., Aurpibul, L., Thompson, P.M., Pruksakaew, K., Lerdlum, S. et al. 2015. Brain imaging and neurodevelopment in HIV-uninfected Thai children born to HIV-infected mothers. *The Pediatric Infectious Disease Journal.* 34(9):e211.

Jankiewicz, M., Holmes, M.J., Taylor, P.A., Cotton, M.F., Laughton, B., Van Der Kouwe, Andre JW & Meintjes, E.M. 2017. White matter abnormalities in children with HIV infection and exposure. *Frontiers in Neuroanatomy.* 11:88.

Jiang, Z., Piggee, C., Heyes, M.P., Murphy, C., Quearry, B., Bauer, M., Zheng, J., Gendelman, H.E. et al. 2001. Glutamate is a mediator of neurotoxicity in secretions of activated HIV-1-infected macrophages. *Journal of Neuroimmunology.* 117(1):97-107.

Jonckheere, H., Anné, J. & De Clercq, E. 2000. The HIV-1 reverse transcription (RT) process as target for RT inhibitors. *Medicinal Research Reviews.* 20(2):129-154.

Joos, B., Fischer, M., Kuster, H., Pillai, S.K., Wong, J.K., Böni, J., Hirschel, B., Weber, R. et al. 2008. HIV rebounds from latently infected cells, rather than from continuing low-level replication. *Proceedings of the National Academy of Sciences.* 105(43):16725-16730.

Kaiser, H.F. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika.* 23(3):187-200.

Kaul, M., Garden, G.A. & Lipton, S.A. 2001. Pathways to neuronal injury and apoptosis in HIV-associated dementia. *Nature.* 410(6831):988.

Keller, M.A., Venkatraman, T.N., Thomas, A., Deveikis, A., LoPresti, C., Hayes, J., Berman, N., Walot, I. et al. 2004. Altered neurometabolite development in HIV-infected children Correlation with neuropsychological tests. *Neurology.* 62(10):1810-1817.

Kerr, S.J., Puthanakit, T., Vibol, U., Aurpibul, L., Vonthanak, S., Kosalaraksa, P., Kanjanavanit, S., Hansudewechakul, R. et al. 2014. Neurodevelopmental outcomes in HIV-exposed-uninfected children versus those not exposed to HIV. *AIDS Care.* 26(11):1327-1335. DOI:10.1080/09540121.2014.920949 [doi].

Kourtis, A.P., Ibegbu, C., Nahmias, A.J., Lee, F.K., Clark, W.S., Sawyer, M.K. & Nesheim, S. 1996. Early progression of disease in HIV-infected infants with thymus dysfunction. *New England Journal of Medicine.* 335(19):1431-1436.

Kurth, A.E., Celum, C., Baeten, J.M., Vermund, S.H. & Wasserheit, J.N. 2011. Combination HIV prevention: significance, challenges, and opportunities. *Current HIV/AIDS Reports.* 8(1):62-72.

Lamers, S.L., Gray, R.R., Salemi, M., Huysentruyt, L.C. & McGrath, M.S. 2011. HIV-1 phylogenetic analysis shows HIV-1 transits through the meninges to brain and peripheral tissues. *Infection, Genetics and Evolution.* 11(1):31-37.

Langford, T.D., Letendre, S.L., Marcotte, T.D., Ellis, R.J., McCutchan, J.A., Grant, I., Mallory, M.E., Hansen, L.A. et al. 2002. Severe, demyelinating leukoencephalopathy in AIDS patients on antiretroviral therapy. *AIDS (London, England).* 16(7):1019.

Laughton, B., Cornell, M., Grove, D., Kidd, M., Springer, P.E., Dobbels, E., van Rensburg, A.J., Violari, A. et al. 2012. Early antiretroviral therapy improves neurodevelopmental outcomes in infants. *AIDS (London, England).* 26(13):1685.

Laughton, B., Cornell, M., Kidd, M., Springer, P.E., Dobbels, E.F.M., Rensburg, A.J.V., Otwombe, K., Babiker, A. et al. 2018. Five year neurodevelopment outcomes of perinatally HIV-infected children on early limited or deferred continuous antiretroviral therapy. *Journal of the International AIDS Society.* 21(5):e25106.

Lentz, M.R., Kim, W.K., Lee, V., Bazner, S., Halpern, E.F., Venna, N., Williams, K., Rosenberg, E.S. et al. 2009. Changes in MRS neuronal markers and T cell phenotypes observed during early HIV infection. *Neurology.* 72(17):1465-1472.

Lewis-de los Angeles, C Paula, Alpert, K.I., Williams, P.L., Malee, K., Huo, Y., Csernansky, J.G., Yogev, R., Van Dyke, R.B. et al. 2016. Deformed subcortical structures are related to past HIV disease severity in youth with perinatally acquired HIV infection. *Journal of the Pediatric Infectious Diseases Society.* 5(suppl_1):S6-S14.

Liauw, L., van der Grond, J., Slooff, V., Wiggers-de Bruine, F., Laan, L., le Cessie, S., van Buchem, M. & van Wezel-Meijler, G. 2008. Differentiation between peritrigonal terminal zones and hypoxic-ischemic white matter injury on MRI. *European Journal of Radiology.* 65(3):395-401.

Lim, K.O., Adalsteinsson, E., Spielman, D., Sullivan, E.V., Rosenbloom, M.J. & Pfefferbaum, A. 1998. Proton magnetic resonance spectroscopic imaging of cortical gray and white matter in schizophrenia. *Archives of General Psychiatry.* 55(4):346-352.

Lin, D.D., Crawford, T.O. & Barker, P.B. 2003. Proton MR spectroscopy in the diagnostic evaluation of suspected mitochondrial disease. *American Journal of Neuroradiology.* 24(1):33-41.

López-Villegas, D., Lenkinski, R.E. & Frank, I. 1997. Biochemical changes in the frontal lobe of HIV-infected individuals detected by magnetic resonance spectroscopy. *Proceedings of the National Academy of Sciences.* 94(18):9854-9859.

Luiz, D.M., Foxcroft, C.D. & Stewart, R. 2001. The construct validity of the Griffiths Scales of Mental Development. *Child: Care, Health and Development.* 27(1):73-83.

MacMaster, F.P. & Kusumakar, V. 2006. Choline in pediatric depression. *McGill Journal of Medicine: MJM.* 9(1):24.

Madhi, S.A., Adrian, P., Cotton, M.F., McIntyre, J.A., Jean-Philippe, P., Meadows, S., Nachman, S., Käyhty, H. et al. 2010. Effect of HIV infection status and anti-retroviral treatment on quantitative and qualitative antibody responses to pneumococcal conjugate vaccine in infants. *The Journal of Infectious Diseases.* 202(3):355-361.

Mahalingam, M., Peakman, M., Davies, E.T., Pozniak, A., McManus, T.J. & Vergani, D. 1993. T cell activation and disease severity in HIV infection. *Clinical and Experimental Immunology.* 93(3):337.

Malee, K.M., Tassiopoulos, K., Huo, Y., Siberry, G., Williams, P.L., Hazra, R., Smith, R.A., Allison, S.M. et al. 2011. Mental health functioning among children and adolescents with perinatal HIV infection and perinatal HIV exposure. *AIDS Care.* 23(12):1533-1544.

Martin, S.C., Wolters, P.L., Toledo-Tamula, M.A., Zeichner, S.L., Hazra, R. & Civitello, L. 2006. Cognitive functioning in school-aged children with vertically acquired HIV infection being treated with highly active antiretroviral therapy (HAART). *Developmental Neuropsychology.* 30(2):633-657.

Mbugua, K., Holmes, M., Cotton, M., Ratai, E., Little, F., Hess, A., Dobbels, E., Van der Kouwe, A. et al. 2016. HIV-associated CD4+/CD8+ depletion in infancy is associated with neurometabolic reductions in the basal ganglia at age 5 years despite early antiretroviral therapy. *Aids.* 30(9):1353-1362. DOI:10.1097/QAD.0000000000001082 Available: http://www.ncbi.nlm.nih.gov/pubmed/26959509.

Mellins, C.A., Smith, R., O'Driscoll, P., Magder, L.S., Brouwers, P., Chase, C., Blasini, I., Hittleman, J. et al. 2003. High rates of behavioral problems in perinatally HIV-infected children are not linked to HIV disease. *Pediatrics.* 111(2):384-393.

Mohamed, M.A., Barker, P.B., Skolasky, R.L., Selnes, O.A., Moxley, R.T., Pomper, M.G. & Sacktor, N.C. 2010. Brain metabolism and cognitive impairment in HIV infection: a 3-T magnetic resonance spectroscopy study. *Magnetic Resonance Imaging.* 28(9):1251-1257.

Mohamed, M.A., Lentz, M.R., Lee, V., Halpern, E.F., Sacktor, N., Selnes, O., Barker, P.B. & Pomper, M.G. 2010. Factor analysis of proton MR spectroscopic imaging data in HIV infection: metabolite-derived factors help identify infection and dementia. *Radiology.* 254(2):577-586.

Montgomery, D.C., Peck, E.A. & Vining, G.G. 2012. *Introduction to linear regression analysis.* John Wiley & Sons.

Mulder, J., McKinney, N., Christopherson, C., Sninsky, J., Greenfield, L. & Kwok, S. 1994. Rapid and simple PCR assay for quantitation of human immunodeficiency virus type 1 RNA in plasma: application to acute retroviral infection. *Journal of Clinical Microbiology.* 32(2):292-300.

Myer, L., Essajee, S., Broyles, L.N., Watts, D.H., Lesosky, M., El-Sadr, W.M. & Abrams, E.J. 2017. Pregnant and breastfeeding women: a priority population for HIV viral load monitoring. *PLoS Medicine.* 14(8).

Nagarajan, R., Sarma, M.K., Thomas, M.A., Chang, L., Natha, U., Wright, M., Hayes, J., Nielsen-Saines, K. et al. 2012. Neuropsychological function and cerebral metabolites in HIV-infected youth. *Journal of Neuroimmune Pharmacology.* 7(4):981-990.

Naismith, S.L., Lagopoulos, J., Hermens, D.F., White, D., Duffy, S.L., Robillard, R., Scott, E.M. & Hickie, I.B. 2014. Delayed circadian phase is linked to glutamatergic functions in young people with affective disorders: a proton magnetic resonance spectroscopy study. *BMC Psychiatry.* 14(1):345.

Nkonki, L.L., Doherty, T.M., Hill, Z., Chopra, M., Schaay, N. & Kendall, C. 2007. Missed opportunities for participation in prevention of mother to child transmission programmes: simplicity of nevirapine does not necessarily lead to optimal uptake, a qualitative study. *AIDS Research and Therapy.* 4(1):27.

Nowak, M.A. & McMichael, A.J. 1995. How HIV defeats the immune system. *Scientific American.* 273(2):58-65.

Nwosu, E.C., Robertson, F.C., Holmes, M.J., Cotton, M.F., Dobbels, E., Little, F., Laughton, B., van der Kouwe, A. et al. 2018. Altered brain morphometry in 7-year old HIV-infected children on early ART. *Metabolic Brain Disease.* 33(2):523-535.

Ouyang, Y., Wei, F., Qiao, L., Liu, K., Dong, Y., Guo, X., Wang, S., Pang, L. et al. 2018. Mitochondrial DNA mutations accumulated in HIV-1-infected children who have an excellent virological response when exposed to long-term antiretroviral therapy. *Journal of Antimicrobial Chemotherapy.* 73(11):3114-3121.

Palmer, S., Josefsson, L. & Coffin, J.M. 2011. HIV reservoirs and the possibility of a cure for HIV infection. *Journal of Internal Medicine.* 270(6):550-560.

Parazzini, C., Baldoli, C., Scotti, G. & Triulzi, F. 2002. Terminal zones of myelination: MR evaluation of children aged 20–40 months. *American Journal of Neuroradiology.* 23(10):1669-1673.

Paul, R.H., Ernst, T., Brickman, A.M., Yiannoutsos, C.T., Tate, D.F., Cohen, R.A. & Navia, B.A. 2008. Relative sensitivity of magnetic resonance spectroscopy and quantitative magnetic resonance imaging to cognitive function among nondemented individuals infected with HIV. *Journal of the International Neuropsychological Society.* 14(5):725-733.

Paul, R., Cho, K., Mellins, C., Malee, K., Robbins, R., Kerr, S., Sophonphan, J., Jahanshad, N. et al. 2019. Predicting neurodevelopmental outcomes in children with perinatal HIV using a novel machine learning algorithm. *bioRxiv.* :632273.

Paul, R., Prasitsuebsai, W., Jahanshad, N., Puthanakit, T., Thompson, P., Aurpibul, L., Hansudewechakul, R., Kosalaraksa, P. et al. 2018. Structural neuroimaging and neuropsychologic signatures in children with vertically acquired HIV. *The Pediatric Infectious Disease Journal.* 37(7):662-668.

Pavlakis, S.G., Lu, D., Frank, Y., Bakshi, S., Pahwa, S., Barnett, T.A., Porricolo, M.E., Gould, R.J. et al. 1995. Magnetic resonance spectroscopy in childhood AIDS encephalopathy. *Pediatric Neurology.* 12(4):277-282. DOI:088789949500048K [pii].

Payne, B.A., Wilson, I.J., Hateley, C.A., Horvath, R., Santibanez-Koref, M., Samuels, D.C., Price, D.A. & Chinnery, P.F. 2011. Mitochondrial aging is accelerated by anti-retroviral therapy through the clonal expansion of mtDNA mutations. *Nature Genetics.* 43(8):806.

Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M. & Ho, D.D. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science.* 271(5255):1582-1586.

Persidsky, Y., Zheng, J., Miller, D. & Gendelman, H.E. 2000. Mononuclear phagocytes mediate blood-brain barrier compromise and neuronal injury during HIV-1-associated dementia. *Journal of Leukocyte Biology.* 68(3):413-422.

Pouwels, P.J., Brockmann, K., Kruse, B., Wilken, B., Wick, M., Hanefeld, F. & Frahm, J. 1999. Regional age dependence of human brain metabolites from infancy to adulthood as detected by quantitative localized proton MRS. *Pediatric Research.* 46(4):474.

Prado, P.T., Escorsi-Rosset, S., Cervi, M.C. & Santos, A.C. 2011. Image evaluation of HIV encephalopathy: a multimodal approach using quantitative MR techniques. *Neuroradiology.* 53(11):899.

Proctor, W.G. & Yu, F.C. 1950. The dependence of a nuclear magnetic resonance frequency upon chemical compound. *Physical Review.* 77(5):717.

Provencher, S.W. 2001. Automatic quantitation of localized in vivo1H spectra with LCModel. *NMR in Biomedicine.* 14(4):260-264.

Puthanakit, T., Ananworanich, J., Vonthanak, S., Kosalaraksa, P., Hansudewechakul, R., van der Lugt, J., Kerr, S.J., Kanjanavanit, S. et al. 2013. Cognitive function and neurodevelopmental outcomes in HIV-infected children older than 1 year of age randomized to early versus deferred antiretroviral therapy: the PREDICT neurodevelopmental study. *The Pediatric Infectious Disease Journal.* 32(5):501.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing.* Available: https://www.R-project.org.

Rae, C.D. 2014. A guide to the metabolic pathways and function of metabolites observed in human brain 1 H magnetic resonance spectra. *Neurochemical Research.* 39(1):1-36.

Randall, S.R., Warton, C.M., Holmes, M.J., Cotton, M.F., Laughton, B., van der Kouwe, Andre JW & Meintjes, E.M. 2017. Larger Subcortical Gray Matter Structures and Smaller Corpora Callosa at Age 5 Years in HIV Infected Children on Early ART. *Frontiers in Neuroanatomy.* 11:95.

Reiser, M.F., Semmler, W. & Hricak, H. 2007. *Magnetic resonance tomography.* Springer Science & Business Media.

Richman, D.D. 1992. HIV drug resistance. *AIDS Research and Human Retroviruses.* 8(6):1065-1071.

Robertson, F.C., Holmes, M.J., Cotton, M.F., Dobbels, E., Little, F., Laughton, B., Van Der Kouwe, André JW & Meintjes, E.M. 2018. Perinatal HIV Infection or Exposure is Associated with Low N-acetylaspartate and Glutamate in Basal Ganglia at age 9 but not 7 Years. *Frontiers in Human Neuroscience.* 12:145.

Robertson, K.R., Smurzynski, M., Parsons, T.D., Wu, K., Bosch, R.J., Wu, J., McArthur, J.C., Collier, A.C. et al. 2007. The prevalence and incidence of neurocognitive impairment in the HAART era. *Aids.* 21(14):1915-1921.

Robertson, K., Liner, J. & Meeker, R.B. 2012. Antiretroviral neurotoxicity. *Journal of Neurovirology.* 18(5):388-399.

Robinson, B., Li, Z. & Nath, A. 2007. Nucleoside reverse transcriptase inhibitors and human immunodeficiency virus proteins cause axonal injury in human dorsal root ganglia cultures. *Journal of Neurovirology.* 13(2):160-167.

Salvan, A., Lamoureux, S., Michel, G., Confort-Gouny, S., Cozzone, P.J. & Vion-Dury, J. 1998. Localized proton magnetic resonance spectroscopy of the brain in children infected with human immunodeficiency virus with and without encephalopathy. *Pediatric Research.* 44(5):755.

Sanders, L.M. & Zeisel, S.H. 2007. Choline: dietary requirements and role in brain development. *Nutrition Today.* 42(4):181.

Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M. & Belshaw, R. 2010. Viral mutation rates. *Journal of Virology.* 84(19):9733-9748.

Sarma, M.K., Nagarajan, R., Keller, M.A., Kumar, R., Nielsen-Saines, K., Michalik, D.E., Deville, J., Church, J.A. et al. 2014. Regional brain gray and white matter changes in perinatally HIV-infected adolescents. *NeuroImage: Clinical.* 4:29-34.

Sepkowitz, K.A. 2001. AIDS—the first 20 years. *New England Journal of Medicine.* 344(23):1764-1772.

Shimizu, M., Suzuki, Y., Yamada, K., Ueki, S., Watanabe, M., Igarashi, H. & Nakada, T. 2017. Maturational decrease of glutamate in the human cerebral cortex from childhood to young adulthood: a 1 H-MR spectroscopy study. *Pediatric Research.* 82(5):749-752.

Siliciano, J.D., Kajdas, J., Finzi, D., Quinn, T.C., Chadwick, K., Margolick, J.B., Kovacs, C., Gange, S.J. et al. 2003. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nature Medicine.* 9(6):727-728. DOI:10.1038/nm880 [doi].

Sisk, C.L. & Zehr, J.L. 2005. Pubertal hormones organize the adolescent brain and behavior. *Frontiers in Neuroendocrinology.* 26(3-4):163-174.

Snider, S.A., Margison, K.D., Ghorbani, P., LeBlond, N.D., O'Dwyer, C., Nunes, J.R., Nguyen, T., Xu, H. et al. 2018. Choline transport links macrophage phospholipid metabolism and inflammation. *Journal of Biological Chemistry.* 293(29):11600-11611.

Soares, D.P. & Law, M. 2009. Magnetic resonance spectroscopy of the brain: review of metabolites and clinical applications. *Clinical Radiology.* 64(1):12-21.

Sonnewald, U. 2014. Glutamate synthesis has to be matched by its degradation–where do all the carbons go? *Journal of Neurochemistry.* 131(4):399-406.

Sowell, E.R., Thompson, P.M., Holmes, C.J., Batth, R., Jernigan, T.L. & Toga, A.W. 1999. Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping. *NeuroImage.* 9(6):587-597.

Spudich, S., Gisslen, M., Hagberg, L., Lee, E., Liegler, T., Brew, B., Fuchs, D., Tambussi, G. et al. 2011. Central nervous system immune activation characterizes primary human immunodeficiency virus 1 infection even in participants with minimal cerebrospinal fluid viral burden. *Journal of Infectious Diseases.* 204(5):753-760.

Steen, C., Wilczak, N., Hoogduin, J.M., Koch, M. & De Keyser, J. 2010. Reduced creatine kinase B activity in multiple sclerosis normal appearing white matter. *PLoS One.* 5(5):e10811.

Steingard, R.J., Yurgelun-Todd, D.A., Hennen, J., Moore, J.C., Moore, C.M., Vakili, K., Young, A.D., Katic, A. et al. 2000. Increased orbitofrontal cortex levels of choline in depressed adolescents as detected by in vivo proton magnetic resonance spectroscopy. *Biological Psychiatry.* 48(11):1053-1061.

Stringer, E.M., Ekouevi, D.K., Coetzee, D., Tih, P.M., Creek, T.L., Stinson, K., Giganti, M.J., Welty, T.K. et al. 2010. Coverage of nevirapine-based services to prevent mother-to-child HIV transmission in 4 African countries. *The Journal of the American Medical Association.* 304(3):293-302.

Tachikawa, M., Fukaya, M., Terasaki, T., Ohtsuki, S. & Watanabe, M. 2004. Distinct cellular expressions of creatine synthetic enzyme GAMT and creatine kinases uCK-Mi and CK-B suggest a novel neuron–glial relationship for brain energy homeostasis. *European Journal of Neuroscience.* 20(1):144-160.

Takeuchi, H., Mizuno, T., Zhang, G., Wang, J., Kawanokuchi, J., Kuno, R. & Suzumura, A. 2005. Neuritic beading induced by activated microglia is an early feature of neuronal dysfunction toward neuronal death by inhibition of mitochondrial respiration and axonal transport. *Journal of Biological Chemistry.* 280(11):10444-10454.

Tarassishin, L., Suh, H. & Lee, S.C. 2011. Interferon regulatory factor 3 plays an anti-inflammatory role in microglia by activating the PI3K/Akt pathway. *Journal of Neuroinflammation.* 8(1):187.

Tavazzi, E., Morrison, D., Sullivan, P., Morgello, S. & Fischer, T. 2014. Brain inflammation is a common feature of HIV-infected patients without HIV encephalitis or productive brain infection. *Current HIV Research.* 12(2):97-110.

Tayebati, S.K., Marucci, G., Santinelli, C., Buccioni, M. & Amenta, F. 2015. Choline-containing phospholipids: structure-activity relationships versus therapeutic applications. *Current Medicinal Chemistry.* 22(38):4328-4340.

Thompson, K.A., Cherry, C.L., Bell, J.E. & McLean, C.A. 2011. Brain cell reservoirs of latent virus in presymptomatic HIV-infected individuals. *The American Journal of Pathology.* 179(4):1623-1629.

Thored, P., Heldmann, U., Gomes-Leal, W., Gisler, R., Darsalia, V., Taneera, J., Nygren, J.M., Jacobsen, S.W. et al. 2009. Long-term accumulation of microglia with proneurogenic phenotype concomitant with persistent neurogenesis in adult subventricular zone after stroke. *Glia.* 57(8):835-849.

Toborek, M., Lee, Y.W., Pu, H., Malecki, A., Flora, G., Garrido, R., Hennig, B., Bauer, H. et al. 2003. HIV-Tat protein induces oxidative and inflammatory pathways in brain endothelium. *Journal of Neurochemistry.* 84(1):169-179.

Toich, J.T., Taylor, P.A., Holmes, M.J., Gohel, S., Cotton, M.F., Dobbels, E., Laughton, B., Little, F. et al. 2018. Functional Connectivity Alterations between Networks and Associations with Infant Immune Health within Networks in HIV Infected Children on Early Treatment: A Study at 7 Years. *Frontiers in Human Neuroscience.* 11:635.

UNAIDS. 2019. *UNAIDS data 2019.* Available: https://www.unaids.org/en/resources/documents/2019/2019-UNAIDS-data [Aug 9, 2019].

Valcour, V., Chalermchai, T., Sailasuta, N., Marovich, M., Lerdlum, S., Suttichom, D., Suwanwela, N.C., Jagodzinski, L. et al. 2012. Central nervous system viral invasion and inflammation during acute HIV infection. *The Journal of Infectious Diseases.* 206(2):275-282.

Vallat, A., De Girolami, U., He, J., Mhashilkar, A., Marasco, W., Shi, B., Gray, F., Bell, J. et al. 1998. Localization of HIV-1 co-receptors CCR5 and CXCR4 in the brain of children with AIDS. *The American Journal of Pathology.* 152(1):167.

Van Dalen, Y.W., Blokhuis, C., Cohen, S., Ter Stege, J.A., Teunissen, C.E., Kuhle, J., Kootstra, N.A., Scherpbier, H.J. et al. 2016. Neurometabolite alterations associated with cognitive performance in perinatally HIV-infected children. *Medicine.* 95(12).

van der Kouwe, André JW, Benner, T., Salat, D.H. & Fischl, B. 2008. Brain morphometry with multiecho MPRAGE. *NeuroImage.* 40(2):559-569.

Van Rie, A., Mupuala, A. & Dow, A. 2008. Impact of the HIV/AIDS epidemic on the neurodevelopment of preschool-aged children in Kinshasa, Democratic Republic of the Congo. *Pediatrics.* 122(1):e123.

Varatharajan, L. & Thomas, S.A. 2009. The transport of anti-HIV drugs across blood–CNS interfaces: summary of current knowledge and recommendations for further research. *Antiviral Research.* 82(2):A99-A109.

Violari, A., Cotton, M.F., Gibb, D.M., Babiker, A.G., Steyn, J., Madhi, S.A., Jean-Philippe, P. & McIntyre, J.A. 2008. Early antiretroviral therapy and mortality among HIV-infected infants. *New England Journal of Medicine.* 359(21):2233-2244.

Waddell, K.W., Zanjanipour, P., Pradhan, S., Xu, L., Welch, E.B., Joers, J.M., Martin, P.R., Avison, M.J. et al. 2011. Anterior cingulate and cerebellar GABA and Glu correlations measured by 1H J-difference spectroscopy. *Magnetic Resonance Imaging.* 29(1):19-24.

Walker, S.P., Wachs, T.D., Gardner, J.M., Lozoff, B., Wasserman, G.A., Pollitt, E., Carter, J.A. & International Child Development Steering Group. 2007. Child development: risk factors for adverse outcomes in developing countries. *The Lancet.* 369(9556):145-157.

Warszawski, J., Tubiana, R., Le Chenadec, J., Blanche, S., Teglas, J., Dollfus, C., Faye, A., Burgard, M. et al. 2008. Mother-to-child HIV transmission despite antiretroviral therapy in the ANRS French Perinatal Cohort. *Aids.* 22(2):289-299.

Weiss, R.A. 1993. How does HIV cause AIDS? *Science.* 260(5112):1273-1279.

WHO. 2006. *WHO | Antiretroviral therapy of HIV infection in infants and children.* Available: http://www.who.int/hiv/pub/paediatric/infants/en/ [Feb 14, 2018].

WHO. 2013. *WHO | 7.1.4 When to start ART in children.* Available: http://www.who.int/hiv/pub/guidelines/arv2013/art/statartchildren/en/ [Jan 31, 2018].

WHO. 2017. *WHO | Guidelines for managing advanced HIV disease and rapid initiation of antiretroviral therapy.* Available: http://www.who.int/hiv/pub/guidelines/advanced-HIV-disease/en/ [Feb 1, 2018].

Wichmann, T. & DeLong, M.R. 1996. Functional and pathophysiological models of the basal ganglia. *Current Opinion in Neurobiology.* 6(6):751-758.

Wiguna, T., Guerrero, A.P., Wibisono, S. & Sastroasmoro, S. 2012. Effect of 12-week administration of 20-mg long-acting methylphenidate on Glu/Cr, NAA/Cr, Cho/Cr, and mI/Cr ratios in the prefrontal cortices of school-age children in Indonesia: a study using 1H magnetic resonance spectroscopy (MRS). *Clinical Neuropharmacology.* 35(2):81-85.

Wilkinson, I.D., Lunn, S., Miszkiel, K.A., Miller, R.F., Paley, M.N., Williams, I., Chinn, R.J., Hall-Craggs, M.A. et al. 1997. Proton MRS and quantitative MRI assessment of the short term neurological response to antiretroviral therapy in AIDS. *Journal of Neurology, Neurosurgery & Psychiatry.* 63(4):477-482.

Williams, K., Westmoreland, S., Greco, J., Ratai, E., Lentz, M., Kim, W., Fuller, R.A., Kim, J.P. et al. 2005. Magnetic resonance spectroscopy reveals that activated monocytes contribute to neuronal injury in SIV neuroAIDS. *The Journal of Clinical Investigation.* 115(9):2534-2545.

Winston, A., Duncombe, C., Li, P.C., Gill, J.M., Kerr, S.J., Puls, R.L., Taylor-Robinson, S.D., Emery, S. et al. 2012. Two patterns of cerebral metabolite abnormalities are detected on proton magnetic resonance spectroscopy in HIV-infected subjects commencing antiretroviral therapy. *Neuroradiology.* 54(12):1331-1339.

Yadav, S.K., Gupta, R.K., Garg, R.K., Venkatesh, V., Gupta, P.K., Singh, A.K., Hashem, S., Al-Sulaiti, A. et al. 2017. Altered structural brain changes and neurocognitive performance in pediatric HIV. *NeuroImage: Clinical.* 14:316-322.

Yiannoutsos, C.T., Ernst, T., Chang, L., Lee, P.L., Richards, T., Marra, C.M., Meyerhoff, D.J., Jarvik, J.G. et al. 2004. Regional patterns of brain metabolites in AIDS dementia complex. *NeuroImage.* 23(3):928-935.

Yongbi, N.M., Payne, G.S., Collins, D.J. & Leach, M.O. 1995. Quantification of signal selection efficiency, extra volume suppression and contamination for ISIS, STEAM and PRESS localized 1H NMR spectroscopy using an EEC localization test object. *Physics in Medicine & Biology.* 40(7):1293.

Zeisel, S.H. & Da Costa, K. 2009. Choline: an essential nutrient for public health. *Nutrition Reviews.* 67(11):615-623.

7. Appendix A: Supplementary results

A.1 Regression analysis models

A.1.1 Linear regression analysis models:

$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5$, weight=(1/Z)

where Y=metabolite concentration, X1=HIV status (HIV+ or HEU vs HU), X2=Age at scan, X3=Sex, X4=ethnicity, X5=gray/white matter content, Z= standard deviation for metabolite concentration.

For analysis of the relationships between metabolites and clinical measures in HIV+ children, X1 is the clinical measure.

A.1.2 Logistic regression analysis models:

$\text{logit}(P) = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6 + \beta_7X_7 + \beta_8X_8$

where P=Probability of HIV infection, X1=Inflammatory factor, X2=PWM axonal factor, X3=Age at scan, X4=Sex, X5=ethnicity, X6=BG gray matter content, X7=MFGM gray matter content and X8= PWM white matter content.

A.2 Comparing metabolite levels between HIV+ and HIV- children

Supplementary table 1: Results of linear regression analysis comparing absolute metabolite concentrations between HIV+ and HIV- children, when adjusting for age at scan, sex, ethnicity and GM or WM content.

		HIV+	
	B	SE	p
Basal Ganglia (N=133: 75 HIV+, 58 HIV-)			
tNAA	0.01	0.13	0.94
NAA	-0.11	0.12	0.35
Glx	0.07	0.23	0.77
Glu	0.06	0.16	0.76
tCho	0.08	0.03	0.002
Ins	0.19[a]	0.11	0.08
tCr	-0.01	0.10	0.94
Cr	0.10	0.15	0.52
Midfrontal gray matter (N=132: 73 HIV+, 59 HIV-)			
tNAA	0.21	0.12	0.09
NAA	0.19	0.12	0.14
Glx	0.04	0.25	0.88
Glu	0.09	0.19	0.64
tCho	0.14	0.03	<0.001
Ins	0.39	0.11	<0.001
tCr	0.08[b]	0.07	0.30
Cr	0.22	0.10	0.03
Peritrigonal white matter (N=129: 73 HIV+, 56 HIV-)			
tNAA	-0.13	0.14	0.37
NAA	-0.10	0.14	0.44
Glx	-0.35[c]	0.21	0.09
Glu	-0.27	0.16	0.09
tCho	0.03	0.03	0.26
Ins	-0.12	0.11	0.24
tCr	-0.12	0.06	0.06
Cr	0.00	0.07	0.98

Unstandardised coefficient (B); Standard error (SE) and p-value (p).

[a] 1 HIV+ outlier excluded; [b] 2 HIV- (HEU) and 4 HIV+ outliers excluded; [c] 1 HIV- (HU) outlier excluded.

A.3 Clinical measure associations with metabolite levels

Supplementary table 2: Relationships between absolute metabolite concentrations in the basal ganglia, and clinical measures.

	CD4 count enrollment			CD4 percent enrollment			CD4 count scan *		
	B	SE	p	B	SE	p	B	SE	p
tNAA	0.0001	0.0001	0.39	-0.0045	0.0070	0.52	-0.0004	0.0002	0.07
NAA	0.0000	0.0001	0.73	-0.0046	0.0062	0.46	-0.0001	0.0002	0.45
Glx	-0.0001	0.0002	0.59	0.0088	0.0168	0.60	-0.0005	0.0005	0.34
Glu	0.0000	0.0001	0.88	0.0091	0.0109	0.41	0.0000	0.0003	1.00
tCho	0.0000	0.0000	0.77	0.0001	0.0018	0.97	0.0000	0.0001	0.54
Ins	0.0000	0.0001	0.88	-0.0029	0.0076	0.70	-0.0007[a]	0.0003	0.02
tCr	0.0000	0.0001	0.93	0.0067	0.0059	0.26	0.0001	0.0002	0.44
Cr	0.0001	0.0001	0.48	-0.0007	0.0103	0.94	-0.0002	0.0003	0.48
	CD4 percent scan *			CD8 count enrollment ***			CD4/CD8 at enrollment ***		
	B	SE	p	B	SE	p	B	SE	p
tNAA	-0.0092	0.0105	0.39	0.0000	0.0001	0.93	-0.0331	0.1108	0.77
NAA	-0.0036	0.0095	0.71	0.0000	0.0001	0.93	-0.0637	0.0976	0.52
Glx	0.0061	0.0258	0.82	-0.0004	0.0002	0.06	0.2712	0.2541	0.29
Glu	0.0054	0.0168	0.75	-0.0003	0.0001	0.06	0.3137[e]	0.1546	0.05
tCho	-0.0013	0.0028	0.63	0.0000[c]	0.0000	0.34	0.0208	0.0265	0.44
Ins	-0.0024	0.0111	0.83	-0.0001[d]	0.0001	0.17	0.0605	0.0944	0.52
tCr	0.0123[b]	0.0084	0.15	-0.0001	0.0001	0.14	0.0532	0.0810	0.51
Cr	-0.0123	0.0151	0.42	-0.0001	0.0001	0.33	0.2066	0.1496	0.17
	Age at ART initiation			Interrupted ART					
	B	SE	p	B	SE	p			
tNAA	-0.0006	0.0007	0.39	-0.1761	0.1373	0.20			
NAA	-0.0003	0.0006	0.68	-0.0481	0.1222	0.70			
Glx	0.0007	0.0018	0.71	-0.0910	0.3355	0.79			
Glu	0.0011	0.0011	0.33	0.2287	0.2150	0.29			
tCho	-0.0001	0.0002	0.70	-0.0607[f]	0.0312	0.06			
Ins	-0.0002	0.0008	0.81	-0.1423	0.1532	0.36			
tCr	0.0006	0.0006	0.33	-0.1679	0.1174	0.16			
Cr	0.0005	0.0010	0.65	-0.2505	0.2030	0.22			

* CD8 data at enrollment, obtained after 12 weeks for 10 participants, was excluded; data was missing for 3 participants.

** CD4 data for 8 participants was >6 months from scan and thus, excluded.

*** VL data at scan for 5 participants was >6 months from scan and thus, excluded.

[a] 6 HIV+ excluded; [b] 2 HIV+, 1 HEU and 2 HU excluded; [c] 4 HIV+, 1 HEU and 1 HU excluded; [d]4 HIV+ excluded; [e] 4 HIV+, 1 HEU and 1 HU excluded; [f]1 HIV+, 1 HEU, 1 HU excluded.

Supplementary table 3: Relationships between midfrontal gray matter metabolite concentrations and clinical measures of disease severity.

	CD4 count enrollment			CD4 percent enrollment			CD4 count scan *		
	B	SE	p	B	SE	p	B	SE	p
tNAA	0.0001	0.0001	0.44	-0.0051	0.0078	0.51	-0.0003	0.0002	0.20
NAA	0.0000	0.0001	0.78	-0.0031	0.0078	0.69	-0.0003	0.0002	0.19
Glx	0.0001	0.0002	0.67	-0.0211	0.0167	0.21	-0.0009	0.0005	0.06
Glu	0.0001	0.0001	0.31	-0.0136	0.0119	0.26	-0.0007[b]	0.0005	0.15
tCho	0.0000	0.0000	0.83	-0.0017	0.0021	0.42	-0.0001	0.0001	0.01
Ins	0.0001	0.0001	0.51	-0.0017	0.0069	0.80	-0.0004	0.0002	0.03
tCr	-0.0001[a]	0.0001	0.01	-0.0039	0.0064	0.54	-0.0003	0.0002	0.11
Cr	0.0001	0.0001	0.39	-0.0043	0.0070	0.54	-0.0003	0.0002	0.20
	CD4 percent scan *			CD8 count enrollment ***			CD4/CD8 at enrollment ***		
	B	SE	p	B	SE	p	B	SE	p
tNAA	-0.0103	0.0114	0.37	0.0001	0.0001	0.44	0.0030	0.1085	0.98
NAA	-0.0103	0.0116	0.38	0.0000	0.0001	0.64	-0.0396	0.0937	0.67
Glx	-0.0283	0.0258	0.28	0.0004	0.0002	0.08	-0.2628	0.2652	0.33
Glu	-0.0210	0.0184	0.26	0.0002	0.0002	0.13	-0.0633	0.1883	0.74
tCho	-0.0047	0.0031	0.14	0.0000	0.0000	0.25	-0.0219	0.0292	0.46
Ins	-0.0177	0.0101	0.09	0.0001	0.0001	0.26	0.0738	0.0952	0.44
tCr	-0.0006	0.0095	0.95	0.0001	0.0001	0.26	-0.1162	0.0848	0.18
Cr	-0.0087	0.0104	0.41	0.0001	0.0001	0.47	0.0005	0.1170	1.00
	Age at ART initiation			Interrupted ART					
	B	SE	p	B	SE	p			
tNAA	-0.0008	0.0008	0.29	-0.0402	0.1558	0.80			
NAA	-0.0002	0.0008	0.79	-0.0476	0.1560	0.76			
Glx	-0.0027	0.0018	0.13	0.1511	0.3399	0.66			
Glu	-0.0016	0.0012	0.21	0.1030	0.2401	0.67			
tCho	-0.0001	0.0002	0.58	-0.0389	0.0415	0.35			
Ins	-0.0007	0.0007	0.30	-0.1089	0.1391	0.44			
tCr	-0.0001	0.0007	0.90	-0.0434	0.1297	0.74			
Cr	0.0007	0.0007	0.31	0.2084	0.1402	0.14			

* CD8 data at enrollment obtained after 12 weeks for 10 participants, was excluded; data missing for 3 participants.

** CD4 data for 8 participants was >6 months from scan and thus, excluded.

*** VL data at scan for 5 participants was >6 months from scan and thus, excluded.

[a] 5 HIV+, 2 HEU excluded; [b] 6 HIV+, 1 HEU, 1 HU excluded.

Supplementary table 4: Relationships between clinical measures and metabolite concentrations in the peritrigonal white matter.

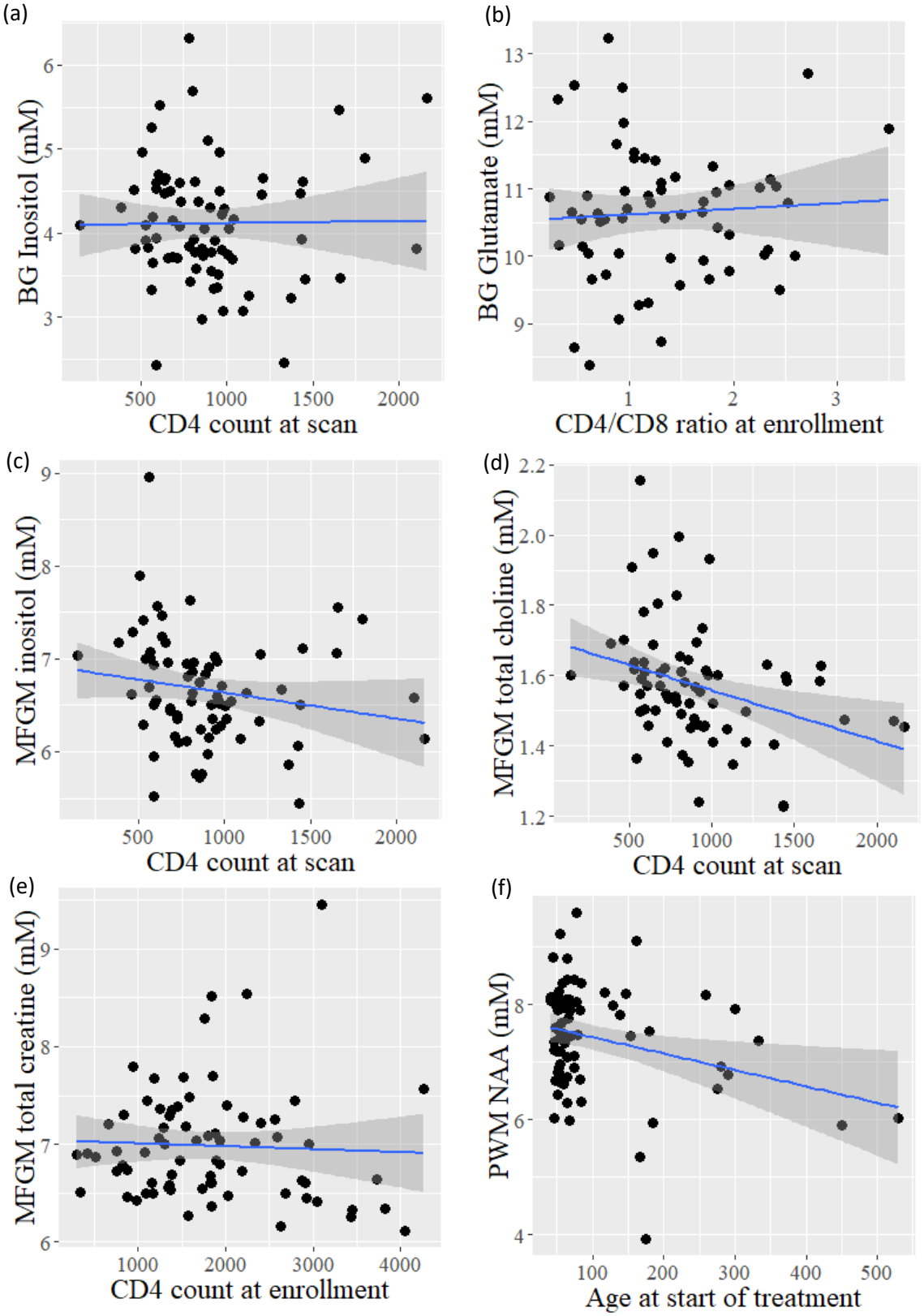
	CD4 count enrollment			CD4 percent enrollment			CD4 count scan *		
	B	SE	p	B	SE	p	B	SE	p
tNAA	0.0001	0.0001	0.38	-0.0050	0.0097	0.61	0.0002	0.0003	0.53
NAA	0.0000	0.0001	0.65	-0.0050	0.0098	0.61	0.0001	0.0003	0.78
Glx	0.0001	0.0002	0.72	-0.0071	0.0151	0.64	0.0002	0.0004	0.70
Glu	0.0001	0.0001	0.52	0.0011	0.0112	0.92	0.0003	0.0003	0.25
tCho	0.0000	0.0000	0.80	-0.0004	0.0020	0.82	0.0000	0.0001	0.61
Ins	0.0000	0.0001	0.71	0.0012	0.0071	0.86	-0.0001	0.0002	0.75
tCr	0.0000	0.0000	0.95	-0.0031	0.0041	0.45	0.0001	0.0001	0.57
Cr	-0.0001	0.0001	0.11	0.0007	0.0047	0.87	0.0000	0.0001	0.86
	CD4 percent scan *			CD8 count enrollment ***			CD4/CD8 at enrollment ***		
	B	SE	p	B	SE	p	B	SE	p
tNAA	0.0107	0.0141	0.45	0.0001	0.0001	0.66	0.0560	0.1485	0.71
NAA	0.0069	0.0141	0.63	0.0001	0.0001	0.61	0.0630	0.1447	0.67
Glx	0.0208	0.0225	0.36	0.0000	0.0002	0.98	0.0625	0.2336	0.79
Glu	0.0189	0.0166	0.26	0.0000	0.0001	0.95	0.0343	0.1644	0.84
tCho	-0.0007	0.0030	0.81	0.0000	0.0000	0.29	0.0286	0.0308	0.36
Ins	-0.0061	0.0106	0.56	-0.0001	0.0001	0.28	0.1751	0.0976	0.08
tCr	0.0046	0.0063	0.46	0.0000	0.0000	0.64	0.0002	0.0554	1.00
Cr	0.0096	0.0052	0.07	0.0000	0.0001	0.89	-0.0268	0.0656	0.69
	Age at ART initiation			Interrupted ART					
	B	SE	p	B	SE	p			
tNAA	-0.0017[a]	0.0009	0.06	-0.1035	0.1953	0.60			
NAA	-0.0022	0.0010	0.03	-0.1331	0.1943	0.50			
Glx	-0.0028	0.0016	0.08	0.0023	0.3037	0.99			
Glu	-0.0021	0.0012	0.07	0.0385	0.2229	0.86			
tCho	-0.0002	0.0002	0.28	-0.0331	0.0402	0.41			
Ins	-0.0008	0.0007	0.26	-0.1704	0.1436	0.24			
tCr	-0.0006	0.0004	0.18	-0.0824	0.0806	0.31			
Cr	0.0002	0.0005	0.66	0.0733	0.0916	0.43			

* CD8 data at enrollment obtained after 12 weeks for 10 participants, was excluded; data missing for 3 participants.

** CD4 data for 8 participants was >6 months from scan and thus, excluded.

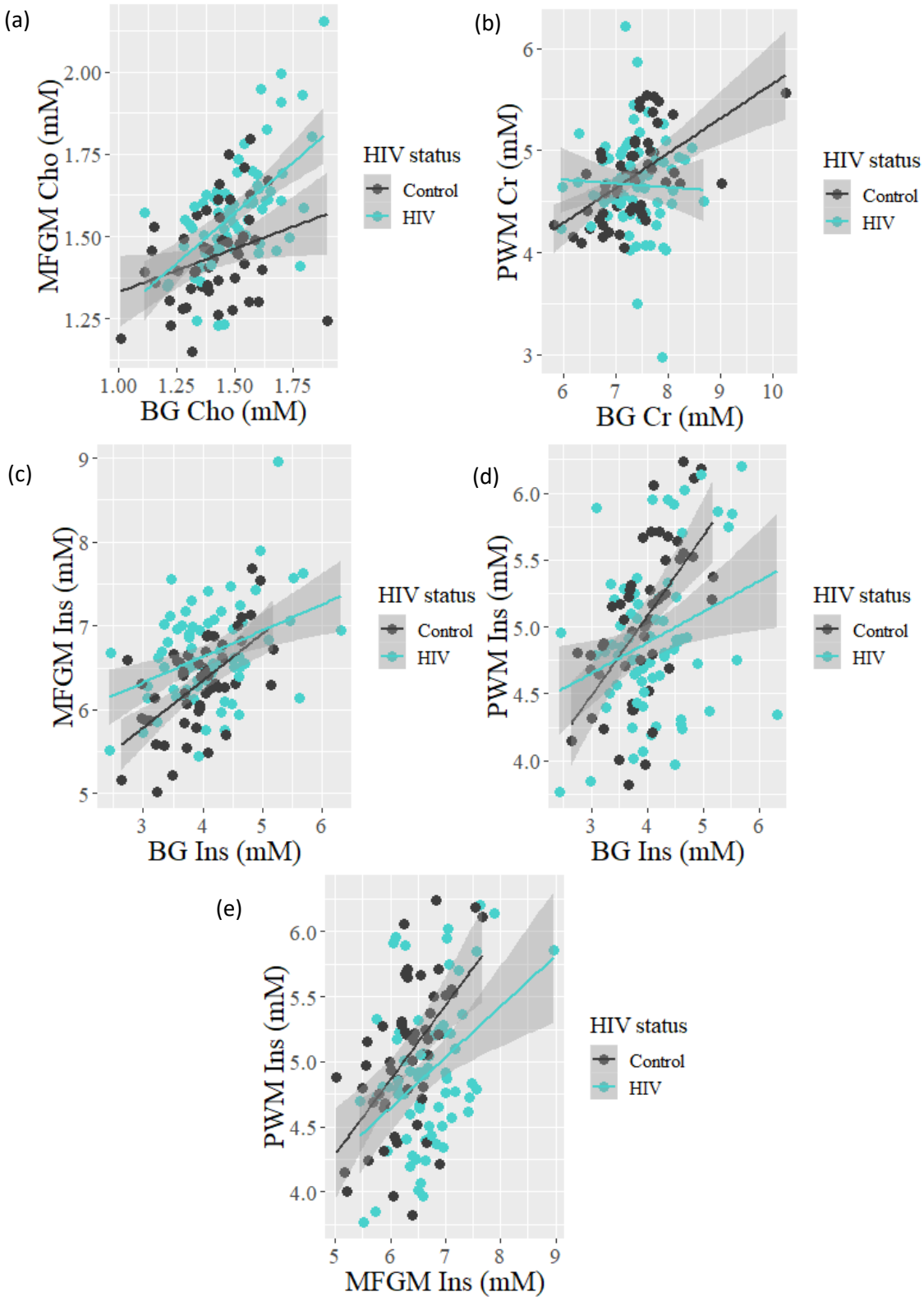
*** VL data at scan for 5 participants was >6 months from scan and thus, excluded.

[a] 3 HIV+ children excluded.



Supplementary figure 1: Plots showing associations that were found between metabolites in the basal ganglia (BG), midfrontal gray matter (MFGM) and peritrigonal white matter (PWM) regions, and clinical measures at enrollment or at scan, specifically for HIV+ children.

A.4 Additional Pearson's correlation results



Supplementary figure 2: Cross-regional metabolite correlations of interest according to whether children are HIV+ or HIV-.

A.5 Additional factor analysis results

Supplementary table 5: Factor loadings for factor analysis with *varimax* rotation, including all 5 metabolites across the 3 regions of interest.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
tCho BG	**0.73**	0.48	<0.1	0.17	<0.1	<0.1	-0.12	0.43
tNAA BG	**0.48**	<0.1	<0.1	0.14	-0.15	<0.1	0.19	0.17
tCr BG	**0.87**	<0.1	<0.1	<0.1	0.13	<0.1	<0.1	-0.11
Glu BG	**0.76**	<0.1	<0.1	<0.1	<0.1	0.16	0.11	<0.1
Ins BG	**0.54**	0.52	<0.1	<0.1	0.32	<0.1	<0.1	-0.18
tCho MFGM	<0.1	**0.75**	0.26	<0.1	<0.1	<0.1	<0.1	0.22
tNAA MFGM	<0.1	0.12	**0.65**	0.13	<0.1	0.33	0.11	0.24
tCr MFGM	<0.1	0.37	**0.91**	0.10	<0.1	0.11	<0.1	<0.1
Glu MFGM	<0.1	0.16	0.36	0.21	<0.1	**0.85**	<0.1	<0.1
Ins MFGM	0.10	**0.64**	0.36	<0.1	0.28	0.24	0.14	<0.1
tCho PWM	<0.1	0.15	0.18	-0.20	0.12	<0.1	<0.1	**0.57**
tNAA PWM	0.23	<0.1	0.11	**0.51**	<0.1	<0.1	0.81	<0.1
tCr PWM	0.15	<0.1	0.27	**0.79**	0.17	<0.1	0.18	<0.1
Glu PWM	<0.1	0.10	<0.1	**0.75**	<0.1	0.17	0.10	-0.17
Ins PWM	<0.1	0.14	<0.1	0.15	**0.96**	<0.1	<0.1	0.13

Supplementary table 6: Logistic regression analysis results for a model assessing the association between HIV status and the **simple** scores of inflammatory and peritrigonal white matter (PWM) axonal factors, when adjusting for confounding variables (N=123: 71 HIV+, 52 HIV-).

	Odds ratio	95% CI	p- value
PWM axonal factor	0.69	0.48-0.95	0.03
Inflammatory factor	1.56	1.26-1.99	< 0.001
Age at scan	2.60	0.52-13.98	0.25
BG gray matter content	1.03	0.97-1.09	0.34
MFGM gray matter content	0.95	0.85-1.06	0.36
PWM white matter content	0.96	0.91-1.02	0.19
Sex (Male)	0.63	0.27-1.45	0.28
Ethnicity (Afrikaans)	0.40	0.12-1.28	0.13

Supplementary table 7: Results of linear regression analysis comparing simple factor scores in HIV+ or HEU children, to those of HU control children.

	HIV+			HEU		
	B	SE	p	B	SE	p
PWM axonal factor	-0.88	0.33	0.01	-1.12	0.41	0.01
Inflammatory factor	1.14	0.52	0.03	-0.63	0.64	0.33
MFGM neuronal factor	0.14	0.39	0.73	-0.70	0.48	0.15
BG energy factor	-0.21	0.21	0.34	-0.25	0.26	0.35
BG neuronal factor	-0.16	0.25	0.52	-0.36	0.31	0.24

Unstandardised coefficient (B); Standard error (SE) and p-value (p).

8. Appendix B: Code in R

```
#==================================
#11-year-old HIV longitudinal study - basal ganglia analysis
#==================================

#Reading in the BG spectroscopy data, demographics spreadsheet and voxel composition data

bgDF <- read.csv("Final BG concentrations.csv", head = TRUE, sep=";")

demoDF <- read.csv("editedCD4CD8VL_11yr_scandata.csv",head=T)

bg_gwm <- read.csv("bg_gwm.csv", head = TRUE, sep=";")

#============================
# Merging 2 data frames by ID
#============================

# Spectroscopy data and demographics for each subject

total_bg <- merge(bgDF, demoDF, by = "Patient.ID")

# Merging with the dataframe containing the percentage of white and gray matter

total_bg2 <- merge(total_bg, bg_gwm, by = "Patient.ID")

#==================================
# Excluding unwanted metabolite data
#==================================

#Creating a new data set only with wanted metabolites

new_bg <- subset(total_bg2, select=c(1:2, 4:6, 15:20, 27:41, 45:50, 60, 63:74, 105, 108, 111:169,
172:174))

#=================================================
#Excluding data which does not meet quality criteria
#=================================================

#Need to exclude data if FWHM >0.07 and SNR <7.0

new_bg$FWHMSNR <- drop(new_bg$FWHM>0.07|new_bg$SNR<7)

table(new_bg$Patient.ID, new_bg$FWHMSNR)

newbg <- subset(new_bg, new_bg$FWHMSNR=="FALSE")

#================================
#Separating into HIV status groups
#================================

#Grouping HIV status into 3 groups

newbg$newArm <- NA

newbg$newArm[newbg$Arm=="HIV unexposed"] <- "HUU"

newbg$newArm[newbg$Arm=="HIV exposed"] <- "HEU"

newbg$newArm[newbg$Arm=="ART-40W"] <- "HIV_P"
```

```
newbg$newArm[newbg$Arm=="ART-96W"] <- "HIV_P"

newbg$newArm[newbg$Arm=="ART-Def"] <- "HIV_P"

newbg$newArm[newbg$Arm=="ART"] <- "HIV_P"

head(newbg)

#=========================================
#Changing the reference group of newArm to HUU
#=========================================

class(newbg$newArm)

newbg$newArm <- as.factor(newbg$newArm)

newbg <- within(newbg, newArm <- relevel(newArm, ref="HUU"))

#=====================
# Printing out outliers ***For all data together*** Rather than according to HIV status ***
#=====================

#GPC.PCh outliers

newbg$Patient.ID[newbg$GPC.PCh %in% boxplot.stats(newbg$GPC.PCh)$out]

boxplot(newbg$GPC.PCh, main="Conc Total Choline", sub=paste("Outlier rows: ",
boxplot.stats(newbg$GPC.PCh)$out))

table(newbg$Patient.ID[newbg$GPC.PCh %in% boxplot.stats(newbg$GPC.PCh)$out],
newbg$GPC.PCh[newbg$GPC.PCh %in% boxplot.stats(newbg$GPC.PCh)$out])

# Removing outliers

Cho.edit <- subset(newbg, Patient.ID!="117" & Patient.ID!="282"& Patient.ID!="324")

#=======================================
#Creating dot plots according to HIV status
#=======================================

#Reading in the ggplot2 package to do this

library(ggplot2)

library(Hmisc)

windowsFonts()

windowsFonts(Times=windowsFont("TT Times New Roman"))


plot <- ggplot(newbg,aes(x=newArm,y=GPC.PCh, fill=newArm))+

 stat_boxplot(geom ='errorbar') +

 geom_boxplot(fill=c("springgreen3", "mediumpurple1", "royalblue"))+

 geom_point(size=2, colour="black", show.legend=FALSE)+

 xlab("HIV status")+
```

```
 ylab("BG total Choline (mM)")+
 ylim(1, 2.5)+
 scale_x_discrete(labels=c("HIV+", "HEU", "HU"),limits=c("HIV_P","HEU","HUU"))+
 theme(axis.text=element_text(size=14, colour = "black"),
      axis.title=element_text(size=16,face="bold"))
plot
#=========================
#Linear regression models ***Entire cohort***Absolute concentrations
#=========================
test <- lm(GPC.PCh~newArm+Gender+Race+Age.11yr.scan+GM,data=newbg,
weights=(1/newbg$GPC.PCh..SD))
summary(test)
#Removing outliers
testb <- lm(GPC.PCh~newArm+Gender+Race+Age.11yr.scan+GM,data=Cho.edit,
weights=(1/Cho.edit$GPC.PCh..SD))
summary(testb)
#Plots to assess the model
par(mfrow=c(2,2))
plot(test)
#=============================================================================
#Analysis of the relationship between CLINICAL MEASURES and metabolite concentrations
#=============================================================================
#Creating a subset containing only the HIV+ children
HIV_P <- subset(newbg, newArm=="HIV_P")
#=========================
#Categorising the viral load data
#=========================
#At enrollment
HIV_P$VL.cat.enrol <- HIV_P$viral.load.baseline
HIV_P$VL.cat.enrol[HIV_P$viral.load.baseline<400] <- "suppressed"
HIV_P$VL.cat.enrol[HIV_P$viral.load.baseline>750000] <- "high"
HIV_P$VL.cat.enrol[HIV_P$viral.load.baseline>=400 & HIV_P$viral.load.baseline<=750000] <- "low"
table(HIV_P$VL.cat.enrol)
#At scan
HIV_P$VL.cat.scan <- HIV_P$VL.Copies.11yr.scan
```

```
HIV_P$VL.cat.scan[HIV_P$VL.Copies.11yr.scan<400] <- "suppressed"

HIV_P$VL.cat.scan[HIV_P$VL.Copies.11yr.scan>750000] <- "high"

HIV_P$VL.cat.scan[HIV_P$VL.Copies.11yr.scan>=400 & HIV_P$VL.Copies.11yr.scan<=750000] <-
"low"

table(HIV_P$VL.cat.scan)

#==================================
#Setting a threshold for clinical data at scan – do not want data > 6 months from scan date (183 days)
#==================================

HIV_P$cd4_threshold <- (abs(HIV_P$clinical_scan_difference_cd4)> 183)

table(HIV_P$cd4_threshold)

HIV_P_cd4scan <- subset(HIV_P, HIV_P$cd4_threshold=="FALSE")


HIV_P$cd8_threshold <- (abs(HIV_P$clinical_scan_difference_cd8)> 183)

table(HIV_P$cd8_threshold)

HIV_P_cd8scan <- subset(HIV_P, HIV_P$cd8_threshold=="FALSE")


HIV_P$VL_threshold <- (abs(HIV_P$clinical_scan_difference_VL)> 183)

table(HIV_P$VL_threshold)

HIV_P_VLscan <- subset(HIV_P, HIV_P$VL_threshold=="FALSE")

#================================
# Setting a threshold for enrollment data
#================================

#CD8 enrollment data needs to be from within age of 12 weeks (i.e. 84 days)

HIV_P$cd8_base_threshold <- (abs(HIV_P$Age_cd8enrol_collected..days.)> 84)

table(HIV_P$cd8_base_threshold, exclude = NULL)

HIV_CD8_enrol_threshold<- subset(HIV_P, HIV_P$cd8_base_threshold =="FALSE")

#Calculating CD4/CD8

HIV_CD8_enrol_threshold$cd4_cd8_ratio <-
(HIV_CD8_enrol_threshold$cd4.count.enrol)/(HIV_CD8_enrol_threshold$cd8_enrolment)

#=======================================
#Checking for outliers in the clinical data
#=======================================

#cd4 count at enrolment

HIV_P$Patient.ID[HIV_P$cd4.count.enrol %in% boxplot.stats(HIV_P$cd4.count.enrol)$out]

boxplot(HIV_P$cd4.count.enrol, main="CD4 count at enrolment", sub=paste("Outlier rows: ",
boxplot.stats(HIV_P$cd4.count.enrol)$out))
```

```
table(HIV_P$Patient.ID[HIV_P$cd4.count.enrol %in% boxplot.stats(HIV_P$cd4.count.enrol)$out],
HIV_P$cd4.count.enrol[HIV_P$cd4.count.enrol %in% boxplot.stats(HIV_P$cd4.count.enrol)$out])

cd4count.edit <- subset(HIV_P, Patient.ID!="116" & Patient.ID!="204")

#=====================================================
#Metabolite outliers + clinical outliers at ENROLLMENT
#=====================================================

#Metabolite outliers + cd4 count outliers at ENROLLMENT

Cho.edit1 <- subset(HIV_P, Patient.ID!="117" & Patient.ID!="282" & Patient.ID!="324" &
Patient.ID!="116" & Patient.ID!="204")

#=====================================================================
#Linear regression looking at the correlation between metabolites and clinical measures
#=====================================================================

#Using absolute concentrations of the metabolites

Model1a <- lm(GPC.PCh~cd4.count.enrol+GM+Age.11yr.scan+Gender+Race,data=HIV_P,
weights=(1/HIV_P$GPC.PCh..SD))

summary(Model1a)

#Linear regression when outliers removed

Model1b <- lm(GPC.PCh~cd4.count.enrol+GM+Age.11yr.scan+Gender+Race,data=Cho.edit1,
weights=(1/Cho.edit1$GPC.PCh..SD))

summary(Model1b)

#*************************************************************************

#=====================
#Analysis across the 3 voxels
#=====================

#Similar analysis as shown above, was carried out for all 3 voxels. Dataframes for the 3 voxels were
#then merged to carry out Pearson's correlation analysis and factor analysis.

#============================
#Merging total data for voxels
#============================

#BG + MFGM data

total_vox <- merge(new_bg, new_mfgm, by="Patient.ID")

#BG/MFGM + PWM data

total_vox2 <- merge(total_vox,  new_pwm, by="Patient.ID")

#=============================================================
#Data cleaning and scaling variables ***Absolute metabolite concentrations
#=============================================================

#Cleaning the total_vox2 dataframe - removing data from scans that did not meet quality checks
across the 3 voxels
```

```
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==227)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==10)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==51)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==74)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==119)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==151)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==187)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==191)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==201)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==21)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==239)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==277)
total_vox2 <- subset(total_vox2, !total_vox2$Patient.ID==84)
table(total_vox2$newArm)

#Recoding HIV status as a binary variable
library(dplyr)
total_vox2$status.x = as.factor(total_vox2$status.x)
total_vox2 <- mutate(total_vox2, status.x.bin = ifelse(status.x == "HIV",1,0))
list(total_vox2$Patient.ID==74, total_vox2$GPC.PCh.y)

#Scaling all the metabolites of interest
total_vox2$cho.BG <- scale(total_vox2$GPC.PCh.x, center = T, scale=T)
total_vox2$naa.BG <- scale(total_vox2$NAA.NAAG.x, center = T, scale=T)
total_vox2$cr.BG <- scale(total_vox2$Cr.PCr.x, center = T, scale=T)
total_vox2$cho.MFGM <- scale(total_vox2$GPC.PCh.y, center = T, scale=T)
total_vox2$naa.MFGM <- scale(total_vox2$NAA.NAAG.y, center = T, scale=T)
total_vox2$cr.MFGM <- scale(total_vox2$Cr.PCr.y, center = T, scale=T)
total_vox2$cho.PWM <- scale(total_vox2$GPC.PCh, center = T, scale=T)
total_vox2$naa.PWM <- scale(total_vox2$NAA.NAAG, center = T, scale=T)
total_vox2$cr.PWM <- scale(total_vox2$Cr.PCr, center = T, scale=T)
total_vox2$glu.BG <- scale(total_vox2$Glu.Gln.x, center = T, scale=T)
```

```
total_vox2$ins.BG <- scale(total_vox2$Ins.x, center = T, scale=T)

total_vox2$glu.MFGM <- scale(total_vox2$Glu.Gln.y, center = T, scale=T)

total_vox2$ins.MFGM <- scale(total_vox2$Ins.y, center = T, scale=T)

total_vox2$glu.PWM <- scale(total_vox2$Glu.Gln, center = T, scale=T)

total_vox2$ins.PWM <- scale(total_vox2$Ins, center = T, scale=T)

#========================
#Pearson's correlation analysis
#========================
#Reading in the necessary packages

library(corrplot)

library(Hmisc)

#Correlations between metabolites within regions and across regions for the ENTIRE COHORT

overall_correl <- subset(total_vox2, select=c(289:291,298:299,292:294,300:301,295:297,302:303))

mycor<- rcorr(as.matrix(overall_correl), type="pearson")

mycor$r

mycor$P

#*****************************************************************************
#Separating the data according to whether children were HIV+ or HIV-

#HIV+ children's data

total_vox6 <- total_vox2[ which(total_vox2$status.x=='HIV'),
c(289:291,298:299,292:294,300:301,295:297,302:303)]

#HIV- children's data

total_vox7 <- total_vox2[ which(total_vox2$status.x=='Control'),
c(289:291,298:299,292:294,300:301,295:297,302:303)]

#*****************************************************************************

#Looking at correlations separately for HIV+ and HIV- children

#For HIV+

pear_HIV<- rcorr(as.matrix(total_vox6), type="pearson")

pear_HIV$r

pear_HIV$P

#For HIV-

pear_control <- rcorr(as.matrix(total_vox7), type="pearson")

pear_control$r

pear_control$P
```

```
#=====================================================================
#Factor analysis ***For 3 factors of interest ***Absolute metabolite concentrations
#=====================================================================

#Creating a dataframe just with the metabolite concentrations of interest

total_vox4 <- subset(total_vox2, select=c(289:297))

#Factor analysis using a maximum likelihood approach and varimax rotation

factanal(total_vox4, factors = 5, rotation = "varimax", sort = TRUE)


#*************************************************************************************

#Manual calculation of weighted factor scores was carried out based on factor loadings determined by factor analysis using the factanal function


total_vox2$fact1 <-
((0.332*total_vox2$cho.MFGM)+(0.707*total_vox2$naa.MFGM)+(0.919*total_vox2$cr.MFGM)+(0.1
70*total_vox2$cho.PWM)+(0.124*total_vox2$naa.PWM)+(0.177*total_vox2$cr.PWM))/(0.332+0.70
7+0.919+0.170+0.124+0.177)

total_vox2$fact2 <-
((0.146*total_vox2$cho.BG)+(0.145*total_vox2$naa.BG)+(0.119*total_vox2$cr.BG)+
(0.160*total_vox2$naa.MFGM)+(0.176*total_vox2$cr.MFGM)+ ( -
0.124*total_vox2$cho.PWM)+(0.583*total_vox2$naa.PWM)+(0.975*total_vox2$cr.PWM))/(0.146+0.
145+0.119+0.160+0.176-0.124+0.583+0.975)

total_vox2$fact3 <-
((0.778*total_vox2$cho.BG)+(0.680*total_vox2$cho.MFGM)+(0.181*total_vox2$naa.MFGM)+(0.301
*total_vox2$cr.MFGM)+(0.394*total_vox2$cho.PWM))/(0.778+0.680+0.181+0.301+0.394)

total_vox2$fact4 <-
((0.476*total_vox2$cho.BG)+(0.229*total_vox2$naa.BG)+(0.954*total_vox2$cr.BG)+(-
0.139*total_vox2$naa.MFGM)+(0.153*total_vox2$cr.MFGM)+(0.110*total_vox2$cr.PWM))/(0.476+
0.229+0.954-0.139+0.153+0.110)

total_vox2$fact5 <-
((0.372*total_vox2$cho.BG)+(0.636*total_vox2$naa.BG)+(0.255*total_vox2$cr.BG)+(-
0.113*total_vox2$cho.MFGM)+(0.224*total_vox2$naa.MFGM)+(0.367*total_vox2$naa.PWM))/(0.37
2+0.636+0.255-0.113+0.224+0.367)


total_vox2$facti <- (total_vox2$naa.MFGM)+(total_vox2$cr.MFGM)

total_vox2$factii <- (total_vox2$naa.PWM)+(total_vox2$cr.PWM)

total_vox2$factiii <- (total_vox2$cho.BG)+(total_vox2$cho.MFGM)+(total_vox2$cho.PWM)

total_vox2$factiv <- (total_vox2$cr.BG)

total_vox2$factv <- (total_vox2$naa.BG)
```

```
#=========================================================================
#Creating a logistic regression model and determining OR and CI for this
#=========================================================================

m1b <- glm(status.x.bin ~ Age.11yr.scan.x + GM.x + GM.y + WM+ Gender.x+ Race.x  +fact2+fact3,
family = binomial(link = logit),

      data = total_vox2)

summary(m1b)

confint(m1b, level = 0.95)

exp(coef(m1b))

exp(confint(m1b, level = 0.95))
```